

Two-photon transitions in primordial hydrogen recombination

Christopher M. Hirata*

Caltech M/C 130-33, Pasadena, California 91125, USA

(Received 11 March 2008; published 1 July 2008)

The subject of cosmological hydrogen recombination has received much attention recently because of its importance to predictions for and cosmological constraints from cosmic microwave background observations. While the central role of the two-photon decay $2s \rightarrow 1s$ has been recognized for many decades, high-precision calculations require us to consider two-photon decays from the higher states ns , $nd \rightarrow 1s$ ($n \geq 3$). Simple attempts to include these processes in recombination calculations with an effective two-photon decay coefficient analogous to the $2s$ decay coefficient $\Lambda_{2s} = 8.22 \text{ s}^{-1}$ have suffered from physical problems associated with the existence of kinematically allowed sequences of one-photon decays, e.g. $3d \rightarrow 2p \rightarrow 1s$, that technically also produce two photons. These correspond to resonances in the two-photon spectrum that are optically thick to two-photon absorption, necessitating a radiative transfer calculation. We derive the appropriate equations, develop a numerical code to solve them, and verify the results by finding agreement with analytic approximations to the radiative transfer equation. The related processes of Raman scattering and two-photon recombination are included using similar machinery. Our results show that early in recombination the two-photon decays act to speed up recombination, reducing the free electron abundance by 1.3% relative to the standard calculation at $z = 1300$. However, we find that some photons between $\text{Ly}\alpha$ and $\text{Ly}\beta$ are produced, mainly by $3d \rightarrow 1s$ two-photon decay and $2s \rightarrow 1s$ Raman scattering. At later times, these photons redshift down to $\text{Ly}\alpha$, excite hydrogen atoms, and act to slow recombination. Thus, the free electron abundance is increased by 1.3% relative to the standard calculation at $z = 900$. Our calculation involves a very different physical argument than the recent studies of Wong and Scott and Chluba and Sunyaev, and produces a much larger effect on the ionization history. The implied correction to the cosmic microwave background power spectrum is negligible for the recently released WMAP and ACBAR data, but at Fisher matrix level will be 7σ for Planck.

DOI: [10.1103/PhysRevD.78.023001](https://doi.org/10.1103/PhysRevD.78.023001)

PACS numbers: 98.70.Vc, 98.62.Ra, 32.80.Rm

I. INTRODUCTION

The anisotropies of the cosmic microwave background (CMB) are one of the most important tools for cosmology. The power spectrum from the Wilkinson Microwave Anisotropy Probe (WMAP) satellite has enabled precision measurement of the composition of the high-redshift universe (the baryon/photon and dark matter/baryon ratios), the distance to the surface of last scattering, and the primordial power spectrum (including its spectral index n_s) [1,2]. The first two of these have proven essential to studies of dark energy, e.g. by breaking the approximate $\Omega_m - w$ degeneracy in low-redshift supernova data, while the measurements of n_s have begun to rule out interesting classes of inflationary models. This trend can be expected to continue with the upcoming Planck satellite and several ground- and balloon-based experiments to measure small-scale temperature anisotropies and CMB polarization. In particular, Planck should be able to map the entire CMB sky with signal-to-noise ratio greater than 1 down to scales of $l \sim 1600$, which in principle enables measurement of n_s with an uncertainty of ~ 0.003 [3].

While the raw sensitivity of future CMB projects is spectacular, there are significant challenges for both experiment and theory. It is well known that these projects require removal of instrumental signatures, foregrounds, and secondary anisotropies, and highly accurate beam maps. In contrast, the theory of primary anisotropies is generally regarded as simple: it is linear perturbation theory whose development was completed more than a decade ago, and which can now be solved by public computer codes that agree to within 0.1% [4]. However, the linear perturbation theory calculation requires knowledge of the free-electron density (and hence Thomson opacity) in the unperturbed Universe. This free-electron density is determined by the complicated nonequilibrium physics of recombination, as first noted 40 years ago by Peebles [5] and Zel'dovich *et al.* [6]. For this reason, a great deal of effort in the 1990s was aimed at solving cosmic recombination. This culminated in the RECFAST code by Seager *et al.* [7,8], which with some recent improvements [9] is used to compute the recombination history in most of today's CMB prediction codes. The Seager *et al.* analysis is based on the "multilevel atom" (MLA) method, in which one writes down the occupation probabilities for each level of each atom or ion, and then constructs a set of evolution equa-

*chirata@tapir.caltech.edu

tions that includes radiative transition rates, photoionization and radiative recombination rates, and collision terms. A separate allowance is made for the two-photon transition $\text{H I } 2s \rightarrow 1s$, and one must also include an equation for the matter temperature (including all heating and cooling terms), since at late times this is not in equilibrium with the CMB temperature. The H I and He II Lyman-series lines, and the $\text{He I } n^1P_1^o - 1^1S_0$ lines, become optically thick and are treated using the Sobolev [10] approximation.

In the past several years, there has been a resurgence of interest in the recombination problem, driven primarily by the new CMB experiments [9,11] and also the possibility of measuring the spectral distortion [12–18]. Several groups have added additional physics to the Seager *et al.* MLA. This includes two-photon transitions from higher excited states of $\text{H I } (ns, nd \rightarrow 1s)$ [19–21] and He I [22]; stimulated two-photon decays and two-photon absorption [23,24]; separate consideration of each nl sublevel of H I instead of assuming a statistical distribution of l [25]; semiforbidden and forbidden transitions in He I [20,26,27]; non-Sobolev behavior in lines with significant continuum opacity and partial redistribution [26,28]; and recoil of H I atoms in $\text{Ly}\alpha$ scattering [29].

This paper is concerned with the two-photon transitions during the $\text{H II} \rightarrow \text{H I}$ recombination, with special attention given to the $ns, nd \rightarrow 1s$ decays with $n \geq 3$. The physics of the two-photon process has been well understood since the work of Göppert-Mayer [30], and successively more accurate computations of the $2s$ lifetime have been made in the following years [31–33]. The importance of the $2s \rightarrow 1s$ decay for cosmological recombination was recognized in the 1960s [5,6], and it is typically included in MLA codes by adding a $2s \rightarrow 1s$ decay with a decay rate $\Lambda_{2s} = 8.22 \text{ s}^{-1}$ and a thermal two-photon absorption rate set by the principle of detailed balance. However, cosmologists paid little attention to the two-photon decays from other states until the recent work of Dubrovich and Grachev [19] (hereafter DG05). Despite the major recent advances in this subject [20,21], we still lack a fully self-consistent treatment of the problem.

DG05 attempted to extend the MLA to two-photon transitions from highly excited ($n \geq 3$) levels by defining analogous two-photon rates $\Lambda_{ns,nd}$. Unfortunately, they discovered that this does not work: the matrix element has poles corresponding to the “ $1 + 1$ ” photon decays of the form $ns, nd \rightarrow Np \rightarrow 1s$ with $1 < N < n$. This results in a very fast decay rate for the $n \geq 3$ levels, with most of the decays producing a photon in a Lyman-series resonance. However, since the photons emitted in resonance lines have already been counted in the one-photon treatment, DG05 recognized that to avoid double counting, they should somehow remove the resonance from their two-photon rates. They did this by keeping only one pole in the matrix element, that corresponding to the np intermediate state. This led them to very fast decay rates that

scale as $\Lambda_{ns,nd} \propto n$ for high n and a several percent acceleration of hydrogen recombination. Wong and Scott [20] (hereafter WS07) compared the DG05 result for $n = 3$ to calculations [34,35] that included all nonresonant poles; they found a lower rate than DG05 and rescaled DG05’s $\Lambda_{ns,nd}$ values appropriately. This resulted in a maximum change of 0.4% in the electron abundance. Chluba and Sunyaev [21] (hereafter CS08) argued that one should cut off the two-photon decay rate not by selecting poles but based on the physical criterion of whether one of the photons is immediately reabsorbed in a Lyman-series line. They computed full two-photon spectra and implemented a cutoff based on their calculation of absorption in the red damping wing of $\text{Ly}\alpha$. This procedure has the distinct advantage of being convergent as $n_{\text{max}} \rightarrow \infty$. The results do depend somewhat on the nature of the cutoff ($x_{\Gamma,c}$ in CS08 notation) but generally give changes in the electron abundance of several tenths of a percent.

In this paper, we write down the full system of equations for two-photon transitions, including emission, absorption, and Raman scattering. We then take two approaches to solving these equations. We first present a numerical approach in which the two-photon continuum is discretized and turned into an effective MLA with virtual levels. We then present an analytic approach in which the two-photon transitions are approximated as effective corrections to the $\text{Ly}\alpha$ and $\text{Ly}\beta$ decay rates and added into the standard MLA with no virtual levels. The first approach, like most fully numerical approaches, has the advantage of solving the two-photon equations with accuracy limited only by step sizes (in both frequency and time!) and machine precision. Unfortunately, the results simply come out of the computer and are difficult to interpret. The analytic approach allows one to understand on paper the cancellations that cause the two-photon effect to be finite and independent of cutoffs. It also serves as a check on the fully numerical result. In the future, it may also have some value in “fast” recombination calculations that could be embedded in Markov chains for cosmological parameter estimation.

This paper is *not* intended to be a complete solution of the hydrogen recombination problem. A number of effects that are not considered here such as the very high- n levels, $\text{Ly}\alpha$ diffusion and recoil, and feedback from the helium spectral distortion [16,18,25,29,36–38], may also be important if subpercent accuracy is desired. Rather, the purpose of this paper is to clear up the physical picture of the two-photon transitions and estimate their effect on recombination, as a first step toward a treatment that will ultimately include these other processes as well.

This paper is organized as follows: in Sec. II, we review the multilevel atom method that is traditionally used in recombination studies and the sometimes-used steady-state approximation. In Sec. III, we describe how to graft two-photon transitions onto this framework, the problems that

have occurred in past efforts to define a two-photon decay rate from highly excited states, and the resolution of these problems. We describe the numerical method for solving the problem and its results in Sec. IV, and an analytic approximation in Sec. V. The effect on the CMB power spectrum is assessed in Sec. VI. We conclude in Sec. VII. The appendices cover several technical details: Appendix A considers the early-time matter temperature evolution, and Appendix B computes the two-photon decay, Raman scattering, and two-photon recombination rates.

Throughout this paper, we use a background cosmology with $\Omega_m h^2 = 0.13$, $\Omega_b h^2 = 0.022$, $T_{\text{CMB}} = 2.728$ K, helium mass fraction $Y = 0.24$, and an effective number of massless neutrinos $N_\nu = 3.04$. This implies the derived parameters $f_{\text{He}} = 0.0795$ (He:H ratio by number) and radiation density $\Omega_r h^2 = 4.196 \times 10^{-5}$. (These parameters are no longer up-to-date but should suffice to illustrate the physics and the magnitude of the corrections due to two-photon transitions.) The cosmological constant is neglected, because it has no influence on the recombination era $\Omega_\Lambda(z) \leq 10^{-8}$. Reduced matrix elements are normalized following the convention of Berestetskii *et al.* [39]. We will define a two-photon decay from a highly excited ($n \geq 3$) level to be “sub-Ly α ” if both emitted photons are below the Ly α frequency and “super-Ly α ” if one photon is emitted above the Ly α frequency, in order to avoid repeating these cumbersome descriptions throughout the paper.

II. THE STANDARD MULTILEVEL ATOM

In this section, we revisit the basic equations of the multilevel atom (Sec. II A) and the steady-state approximation (Sec. II B).

Most of the notation of this section is conventional. We define n_{H} to be the total physical density of hydrogen nuclei (units of cm^{-3}) and define the abundance of hydrogen atoms in level i relative to this total $x_i = n_i/n_{\text{H}}$. A similar definition is used for free electrons and protons $x_e = n_e/n_{\text{H}}$ and $x_p = n_p/n_{\text{H}}$. We denote the energy of a bound state by $E_i < 0$ and the energy of a free electron by $E_e > 0$. The degeneracies of states are denoted by g_i ; in the case of hydrogen, where we do not resolve fine structure levels $g_{nl} = 2(2l + 1)$. Einstein coefficients are denoted A_{ji} for a transition from upper level j to lower level i , and transition energies/frequencies are denoted $E_{ji} \equiv E_j - E_i$ and $\nu_{ji} = E_{ji}/h$. The photon phase-space density is $f(E)$. When considering the phase-space density just above (blueward of) or below (redward of) a line, we use the notation

$$f_{ji\pm} = f(E_{ji} \pm \epsilon), \quad (1)$$

where ϵ is the line width, which is assumed to be infinitesimal (i.e. the time required to redshift through the line is assumed to be much less than the duration of recombina-

tion). Temperatures are quoted in energy units, so that Boltzmann’s constant is equal to 1.

A. Basic equations

The evolution equation for x_i is

$$\dot{x}_i = \dot{x}_i|_{\text{bb}} + \dot{x}_i|_{\text{bf}} + \dot{x}_i|_{2\gamma}, \quad (2)$$

where we have separated the net production of level i into terms coming from radiative bound-bound, radiative bound-free, and two-photon processes.

The bound-bound term is

$$\begin{aligned} \dot{x}_i|_{\text{bb}} = & \sum_{j>i} A_{ji} P_{ji} \left[(1 + f_{ji+}) x_j - \frac{g_j}{g_i} f_{ji+} x_i \right] \\ & + \sum_{j<i} A_{ij} P_{ji} \left[\frac{g_i}{g_j} f_{ij+} x_j - (1 + f_{ji+}) x_i \right], \end{aligned} \quad (3)$$

where the sums are over levels j that are above ($j > i$) or below ($j < i$) the energy of level i ; A_{ji} is the Einstein coefficient; g_i and g_j are the level degeneracies; and f_{ji+} is the photon phase-space density on the blue (incoming) side of the line with frequency $\nu_{ji} = (E_j - E_i)/h$, which is necessary for following absorption and stimulated emission. Here P_{ji} is the Sobolev escape probability for the ji line, which is the probability that a photon emitted in the line will escape via redshifting without being reabsorbed. It is given by

$$P_{ji} = \frac{1 - e^{-\tau_{ji}}}{\tau_{ji}} \quad (4)$$

with the optical depth

$$\tau_{ji} = \frac{c^3 n_{\text{H}}}{8\pi H \nu_{ji}^3} A_{ji} \left(\frac{g_j}{g_i} x_i - x_j \right); \quad (5)$$

see e.g. Ref. [8] for a derivation. In practice, the optical depth is $\tau_{ji} \ll 1$ and $P_{ji} \approx 1$ except for the Lyman-series lines. The jump condition for the radiation field across the line is obtained by finding the net rate of radiative de-excitations (i.e. photons emitted in the line), and multiplying by the conversion factor $8\pi H \nu_{ji}^3 / (c^3 n_{\text{H}})$, which is the number of photon modes that redshift through the line per hydrogen nucleus per unit time

$$f_{ji-} - f_{ji+} = \frac{8\pi H \nu_{ji}^3}{c^3 n_{\text{H}}} A_{ji} P_{ji} \left[x_j (1 + f_{ji+}) - \frac{g_j}{g_i} x_i f_{ji+} \right]. \quad (6)$$

The bound-free term includes spontaneous and stimulated recombination, and the inverse process, photoionization (see e.g. Sec. 2.3.1 of Ref. [8]). It is

$$\begin{aligned} \dot{x}_i|_{\text{bf}} = & \int \alpha_i(E_e) \left\{ P_{\text{M}}(E_e) n_{\text{H}} x_e x_p [1 + f(E_e - E_i)] \right. \\ & \left. - \frac{dg}{dE_e dV} \frac{x_i}{g_i} f(E_e - E_i) \right\} dE_e, \end{aligned} \quad (7)$$

where E_e is the energy of the free electron (more accurately, the center-of-mass energy of the electron and proton), $f(E_e - E_i)$ is the photon phase-space density at specified energy, $\alpha_i(E_e)$ is the recombination coefficient to level i for an electron and proton, and $P_M(E_e)$ is the Maxwellian probability distribution for the electron energy at matter temperature T_m

$$P_M(E_e)dE_e = \frac{2}{\sqrt{\pi}} \frac{E_e^{1/2}}{T_m^{3/2}} e^{-E_e/T_m} dE_e. \quad (8)$$

The $1 + f(E_e - E_i)$ factor accounts for stimulated recombination. The factor $dg/(dE_e dV)$, i.e. the number of available electron states per unit volume per unit energy, is required in order to satisfy the principle of detailed balance. It is equal to

$$\frac{dg}{dE_e dV} = \frac{8\sqrt{2}\pi}{h^3} \mu^{3/2} E_e^{1/2}, \quad (9)$$

where μ is the proton-electron reduced mass, $\mu = m_e/(1 + m_e/m_p)$. Thus, we may write

$$\begin{aligned} \dot{x}_i|_{\text{bf}} = & \frac{2}{\sqrt{\pi}} \int \alpha_i(E_e) \frac{E_e^{1/2}}{T_m^{3/2}} \left\{ n_H x_e x_p [1 + f(E_e - E_i)] e^{-E_e/T_m} \right. \\ & \left. - \frac{2(2\pi T_m)^{3/2}}{h^3} \mu^{3/2} E_e^{1/2} \frac{x_i}{g_i} f(E_e - E_i) \right\} dE_e. \end{aligned} \quad (10)$$

Finally the usual two-photon term is

$$\dot{x}_{2s}|_{2\gamma} = -\dot{x}_{1s}|_{2\gamma} = \Lambda_{2s}(-x_{2s} + x_{1s}e^{-E_{2s,1s}/T_r}), \quad (11)$$

where Λ_{2s} is the two-photon decay rate from the $2s$ level, and the second term accounts for detailed balance.

Note that the bound-bound radiative transition rate requires knowledge of the photon phase-space density $f(E_{ji+})$ on the blue side of the line. Seager *et al.*[8] assumed that this is simply a blackbody function

$$f(E_{ji+}) = \frac{1}{e^{E_{ji}/T_m} - 1}. \quad (12)$$

In reality, the phase-space density for the Lyman-series lines will be greater than the blackbody value, because photons from the $\text{Ly}(n+1)$ line will redshift downward and add to the Lyn line [26,40]. In the absence of any processes that operate between the Lyn and $\text{Ly}(n+1)$ lines, one may simply use phase-space conservation along a photon trajectory

$$f_{np,1s+}(z) = f_{(n+1)p,1s-} \left[\frac{1 - (n+1)^{-2}}{1 - n^{-2}} (1+z) - 1 \right]. \quad (13)$$

This equation refers to a photon phase-space density at some earlier time. We generate a lookup table of $f_{(n+1)p,1s-}$ as we proceed through recombination, and when a value is required we obtain it by 3-point quadratic interpolation. This method requires that the step size $\Delta a/a$ be smaller

than the spacing of the Lyman lines $\nu_{(n+1)p,1s}/\nu_{np,1s} - 1$, which imposes a significant but tractable computational burden.

The matter evolution equation is

$$\dot{T}_m = -2HT_m + \frac{8x_e \sigma_T a_r T_r^4}{3(1 + f_{\text{He}} + x_e)m_e c} (T_r - T_m), \quad (14)$$

where the second term comes from electron scattering [8]. It contains the Thomson cross section σ_T and the radiation constant a_r . The second term is very effective at driving the matter temperature toward the radiation temperature at early times. Later, when x_e and T_r decline the matter temperature falls below the radiation temperature. Since the radiation temperature scales as $T_r \propto a^{-1}$, we may write this equation as

$$\frac{d}{dt} \left(\frac{T_m}{T_r} \right) = -H \frac{T_m}{T_r} + \frac{8x_e \sigma_T a_r T_r^4}{3(1 + f_{\text{He}} + x_e)m_e c} \left(1 - \frac{T_m}{T_r} \right). \quad (15)$$

B. Steady-state approximation

We now introduce the steady-state approximation, which is used to accelerate solution of the hydrogen recombination problem by eliminating the need for a stiff differential equation solver.

The above equations for the excited hydrogen levels ($i \neq 1s$) can be written in the form

$$\dot{x}_i = s_i + \sum_{j \neq 1s} (R_{ji}x_j - R_{ij}x_i) - \beta_i x_i - \gamma_i x_i, \quad (16)$$

where s_i is a source term that depends on excitations from $1s \rightarrow i$ and recombinations to level i , and R_{ij} is the transition rate from level i to level j , β_i is the rate of ionizations from level i , and γ_i is the rate of decays (spontaneous, stimulated emission) from level i to the ground state. The coefficients $\{s_i, R_{ij}, \beta_i, \gamma_i\}$ in general depend on the ionization fraction, radiation and matter temperatures, free-electron density, and spectral distortions (since they involve f_{ji+}). The explicit expressions are for the source term

$$\begin{aligned} s_i = & \frac{g_i}{2} A_{i,1s} P_{i,1s} f_{i,1s+} x_{1s} + \frac{2}{\sqrt{\pi}} n_H x_e x_p \int [1 + f(E_e - E_i)] \\ & \times \frac{\alpha_i(E_e) E_e^{1/2}}{T_m^{3/2}} e^{-E_e/T_m} dE_e + x_{1s} \delta_{i,2s} \Lambda_{2s} e^{-E_{2s,1s}/T_r}; \end{aligned} \quad (17)$$

for the transition term,

$$R_{ji} = \begin{cases} A_{ji} P_{ji} (1 + f_{ji+}) & j > i; \\ A_{ij} P_{ij} f_{ij+} (g_i/g_j) & j < i; \end{cases} \quad (18)$$

for the ionization term,

$$\beta_i = \frac{2^{7/2} \pi \mu^{3/2}}{h^3} \int \alpha_i(E_e) \frac{E_e}{g_i} f(E - E_i) dE_e; \quad (19)$$

and for the decay term,

$$\gamma_i = A_{i,1s} P_{i,1s} (1 + f_{i,1s+}) + \delta_{i,2s} \Lambda_{2s}. \quad (20)$$

In these equations, there is a dependence on x_e , but almost no dependence on the x_i s. The exceptions are (i) formally $x_{1s} = 1 - x_e - \sum_i x_i$ depends on the populations of the excited states, and (ii) the Sobolev probabilities P_{ij} contain a small dependence on the excited states. However, past analyses have shown that the populations of the excited states are very small (the maximum is $x_{2p} = 5.5 \times 10^{-14}$ at $z = 1360$), so we may take $x_{1s} \approx 1 - x_e$. Also, the Sobolev escape probabilities for transition between two excited state are always very close to 1 (the peak optical depth for such lines is 5.5×10^{-4} for $H\alpha$ at $z = 1400$, corresponding to an escape probability of 0.99 972). Finally, the Sobolev probability for transitions involving the $1s$ level are very small, but the high optical depth is due entirely to the $1s$ level—the contribution from the excited level np changes the optical depth by a factor of $x_{np}/3x_{1s}$, which has a maximum value of 1.7×10^{-12} (for $n = 2$, $z = 1600$; at later times or for higher n it is lower). This leads to a convenient repackaging of Eq. (16)

$$\dot{x}_i = s_i - \sum_j T_{ij} x_j, \quad (21)$$

where

$$T_{ij} = \delta_{ij} \left(\beta_i + \gamma_i + \sum_k R_{ik} \right) - R_{ji} \quad (22)$$

is a square matrix.

The excited states of a hydrogen atom typically have very short lifetimes compared with the recombination timescale: the longest intrinsic lifetime is that of $2s$ ($\Lambda_{2s}^{-1} = 0.12$ s), and even taking into account Sobolev suppression the intrinsic lifetimes of the p states are short, e.g. for $2p$ we have $(A_{2p,1s} P_{2p,1s})^{-1} \sim 1$ s when the Ly α optical depth reaches its peak of $\sim 6 \times 10^8$. For comparison, hydrogen recombination takes several $\times 10^{12}$ s. Therefore, we expect that to very high accuracy the excited level populations should be in steady state given Eq. (21), i.e. we should have

$$x_i \approx \sum_j (\mathbf{T}^{-1})_{ij} s_j \quad (23)$$

and

$$\dot{x}_e \approx -\dot{x}_{1s} = -\sum_{i>1s} \left(\gamma_i x_i - \frac{g_i}{2} A_{i,1s} P_{i,1s} f_{i,1s+} x_{1s} \right). \quad (24)$$

Formally, it is the minimum eigenvalue t_r^{-1} of the matrix \mathbf{T} that indicates how rapidly a steady-state solution is approached. During the redshift range covered by our code

the relaxation time t_r (reciprocal of the minimum eigenvalue of \mathbf{T}) has a maximum of 0.8 s, or $\sim 10^{-12}$ of the duration of recombination. Thus, the steady-state approximation should very accurately describe the populations of the excited states and the ionization history as a function of time. (It may not accurately describe the low-frequency spectral distortion, since the latter depends on slight deviations of the excited state ratios from thermal equilibrium, i.e. one is taking differences of excited state populations that are nearly zero. In general, the spectral distortion is much more sensitive than the recombination history to numerical errors [12,13].)

In the case of the Lyman lines where there is significant feedback, we need the phase-space density $f_{(n+1)p,1s-}$, which is easily obtained from Eq. (6) once we have obtained $x_{(n+1)p}$ from Eq. (23).

In order to eliminate the need for a stiff ordinary differential equation (ODE) solver we also need to remove the rapidly decaying mode from the matter temperature equation, Eq. (15). At early times, this can be done using the first-order asymptotic solution

$$\frac{T_m}{T_r}(t) = 1 - \frac{3(1 + f_{\text{He}} + x_e)m_e c H}{8x_e \sigma_T a_r T_r^4}. \quad (25)$$

(See Appendix A for a derivation). This equation can be used to solve algebraically for T_m at each step in the integration, rather than requiring a stiff integrator. In fact, it is quite accurate at early times: the RECFAST code, which fully integrates Eq. (15) at low redshifts, gives a value for T_m differing by $<0.01\%$ from Eq. (25) for all $z \geq 700$. (At early times, RECFAST sets $T_m = T_r$, since the Compton equilibrium time is fast, but we have checked that at $z \leq 850$, RECFAST has switched on the integration of the \dot{T}_m equation, so this is a fair comparison.) Accurate calculations at lower redshift would require one to “switch on” the differential equation for T_m once the matter-radiation equilibration time $(HJ)^{-1}$ becomes a significant fraction of the recombination time; at this point, it would be possible to follow T_m with a nonstiff integrator.

In most implementations of the MLA, one first groups states together and then solves either the “full” Eq. (2) or the steady-state problem, Eq. (24). For example, the original calculation of Peebles [5] used a single, lumped excited state and the steady-state approximation (with some additional approximations such as neglect of stimulated recombination).

III. EXTENSION TO TWO-PHOTON TRANSITIONS

Now, we would like to extend our analysis to include two-photon decays from the high-lying levels. The selection rules for two-photon electric dipole decays restrict us to excited levels ns and nd (two-photon $np \rightarrow 1s$ decays are parity-forbidden). At first sight this is a simple problem. The two-photon decay rates can be calculated in tree-level QED, as done in Appendix B, and the total transition

rate Λ_{nl} can be found using Eq. (B6). This rate can be added to the evolution equations in analogy to Eq. (11).

Unfortunately, this program does not work because the desired rates of e.g. spontaneous two-photon decay $3d \rightarrow 1s$ double count the “1 + 1” or cascade decay process $3d \rightarrow 2p \rightarrow 1s$, and do not correct for reabsorption of the emitted photons. We develop our understanding of this issue by first discussing the two-photon decay rates in Sec. III A and then reviewing previous attempts to incorporate two-photon transitions in the MLA (Sec. III B). We discuss the physical resolution of the double counting and reabsorption problems in Sec. III C. We then write down the full radiative transfer equations in Sec. III D. The principal result of this paper will be the solution of this last set of equations.

A. Rate coefficients

The theory and computation of two-photon decay rate coefficients is reviewed in Appendix B. Here, we merely recount the most important facts and results. The two-photon decay is

$$H(nl) \rightarrow H(1s) + h\nu + h\nu', \quad (26)$$

where we define ν to be the higher-energy photon and ν' to be the lower-energy photon, so that there is no issue of indistinguishable particles and the corresponding factors of $1/(2!)$. The two photons have frequencies that sum to $(1 - n^{-2})\mathcal{R}$ due to conservation of energy. The rate of decays is given by an integral of the differential rate

$$\Gamma_{nl \rightarrow 1s} = \int_{(1-n^{-2})\mathcal{R}/2}^{(1-n^{-2})\mathcal{R}} \frac{d\Gamma}{d\nu} d\nu, \quad (27)$$

where in a thermal radiation bath the actual decay rate is related to the spontaneous rate $d\Lambda/d\nu$ by

$$\frac{d\Gamma}{d\nu} = \frac{d\Lambda}{d\nu} (1 - e^{-h\nu/T_r})^{-1} (1 - e^{-h\nu'/T_r})^{-1}. \quad (28)$$

The increase over the vacuum ($T_r = 0$) rate is due to stimulated two-photon decays [23].

Panel (a) of Fig. 1 shows the two-photon decay rate from $2s \rightarrow 1s$ at several temperatures. The differential decay rate $d\Gamma/d\nu$ is well-defined and integrable at all temperatures, and this decay presents no special problems.

Panel (b) of Fig. 1 shows the $3s \rightarrow 1s$ decays. Here, Eq. (28) contains a resonance at $\nu = 3\mathcal{R}/4$, i.e. the Ly α frequency. This corresponds to the pole in the matrix element from the $2p$ intermediate state. Taking the two-photon decay formula at face value the total rate $\Gamma_{3s \rightarrow 1s}$ is infinite due to the $\sim 1/\Delta\nu^2$ divergence in the rate. If one includes the imaginary energy (“pole displacement”) of $2p$ due to its finite lifetime (1.6 ns), then one gets a finite decay rate, but very large: in vacuum $\Gamma_{3s \rightarrow 1s} = 6.3 \times 10^6 \text{ s}^{-1}$. This is equal to the one-photon transition rate $A_{3s,2p}$ (see discussion in CS08) for a simple physical reason: the sequential or “1 + 1” decay $3s \rightarrow 2p \rightarrow 1s$ is simply a two-photon decay with the photons on resonance. (The one- and two-photon contributions can be separated as different diagrams that contribute to the width or imaginary energy shift of the $3s$ level $\Im \Delta E_{3s}$ [41,42], but $\Re \Delta E_{3s}$

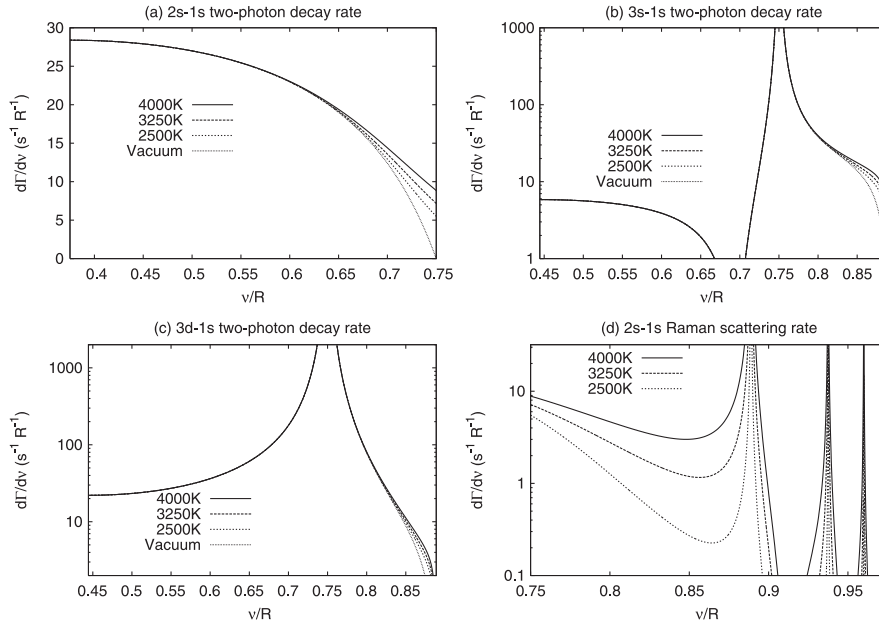


FIG. 1. The two-photon transition rates for the most important excited levels of H I. The horizontal axis shows the frequency of the higher-energy photon (for 2γ decays) or the outgoing photon (for $2s \rightarrow 1s$ Raman scattering). The differential transition rate $d\Gamma/d\nu$ is shown at several different temperatures and in vacuum ($T_r = 0$). The two-photon spectra are shown from half the maximum frequency ($\nu_{\max}/2$) to ν_{\max} and are symmetric. The Ly α resonance at $\nu = 0.75\mathcal{R}$ is easily seen in the $3s$, $3d \rightarrow 1s$ rates.

does not give us the frequency distribution of emitted photons, which we need in our calculations.) A similar problem occurs for the $3d$ initial level, as seen in panel (c) of Fig. 1.

A similar problem occurs for Raman scattering

$$H(nl) + h\nu' \rightarrow H(1s) + h\nu. \quad (29)$$

This time it is the difference of frequencies that is fixed by energy conservation $\nu = \nu' + (1 - n^{-2})\mathcal{R}$. There is no Raman scattering in vacuum ($T_r = 0$) as it requires an initial-state photon, but one can define a Raman-scattering rate in analogy to Eq. (28) at finite temperature. This is shown for the $2s$ initial state in the panel (d) of Fig. 1. In this case, there are resonances when the virtual energy of the intermediate state corresponds with an energy level of hydrogen: at $\nu = 8\mathcal{R}/9$, $15\mathcal{R}/16$, etc., and they reproduce the thermal rate for the sequential $1 + 1$ photon transitions $2s \rightarrow 3p \rightarrow 1s$, $2s \rightarrow 4p \rightarrow 1s$, etc.

For completeness, we also discuss two-photon recombination

$$H^+ + e^- \rightarrow H(1s) + h\nu + h\nu'. \quad (30)$$

The differential cross sections $d\sigma/d\nu$ for this process are shown in Fig. 2. It too exhibits Lyman-series resonances. Note that the differential rate coefficient $\alpha^{(2)}(\nu, \nu')$ is equal to the incident electron velocity times $d\sigma/d\nu$: $\alpha^{(2)}(\nu, \nu') = v_e d\sigma/d\nu$.

At this stage, we have two basic problems. One is the double-counting problem: we have already included $1 + 1$ photon transitions in the MLA, so it seems incorrect to recount the resonant two-photon transitions. The other is the reabsorption problem: the very large two-photon transition rates near resonance imply large optical depth for reabsorption of some of the emitted radiation, combined with the possibility that the Lyman-line spectral distortions

could be immediately reabsorbed. The next section critically discusses previous attempts to solve these problems.

B. Past attempts at inclusion in MLA codes

The double-counting and reabsorption problems were recognized even in the very first attempts to include two-photon decays in the MLA. There have been three major approaches suggested in the recombination literature so far: those of DG05, WS07, and CS08. Here, we describe each of them and explain why, while each of these methods contains important physical insights, none of them are both complete and applicable to the problem at hand. In each of these cases, the authors themselves noted that their solution was only a first approximation and that ultimately a radiative transfer calculation was required; in particular, CS08 provided a detailed description of the problem.

DG05 avoided the double-counting problem by removing the $1 + 1$ poles from the two-photon decay matrix element \mathcal{M} . For example, in the case of the $3d \rightarrow 1s$ decay, they removed the $2p$ intermediate state. They then argued that in $ns, nd \rightarrow 1s$ decays, the largest contribution to the matrix element comes from the np intermediate state. Keeping only this term in \mathcal{M} they arrived at a finite decay rate:

$$\Lambda_{nl}^{\text{DG05}} \propto \frac{\nu_{nl,1s}^5}{2l+1} \left| \langle nl || \mathbf{r} || np \rangle \langle np || \mathbf{r} || 1s \rangle \right|^2, \quad (31)$$

which they normalize to the $2s$ decay rate $\Lambda_{2s} = 8.22 \text{ s}^{-1}$. They then substitute this rate coefficient into their MLA code in the same way that the traditional treatment inserts Λ_{2s} . The DG05 rate coefficients scale as $\propto n$, and so a cutoff must be imposed; they impose such a cutoff when the wavelength of emitted radiation is less than the size of the atom ($\sim a_0 n^2$), because at higher n the electric dipole formula is no longer valid.

This approach obviously eliminates the double-counting problem. Also, by getting rid of the large differential rate $d\Lambda/d\nu$ near resonance it eliminates the reabsorption problem. However, it ignores the destructive interference in \mathcal{M} that arises from consideration of the different intermediate states; it predicts $d\Lambda/d\nu \propto n$ for large n and fixed ν , whereas the actual scaling when interference is considered is $\propto n^{-3}$. This occurs because while the largest contribution to the $ns, nd \rightarrow 1s$ matrix element comes from the np intermediate state, the neighboring intermediate states $(n \pm 1)p$, $(n \pm 2)p$, etc. give opposite sign contributions and result in a cancellation of the matrix element that becomes more exact as $n \rightarrow \infty$ [22]. Thus, to have even a qualitatively correct result for $d\Lambda/d\nu$ we must include the troublesome intermediate states, such as $(n-1)p$, even if they are far from resonance.

WS07 presented an improved treatment of the problem. They used rate coefficients for $3s \rightarrow 1s$ and $3d \rightarrow 1s$ two-photon decays that included all the $n \geq 3$ intermediate states (i.e. those that cannot be on shell) [34,35]. From

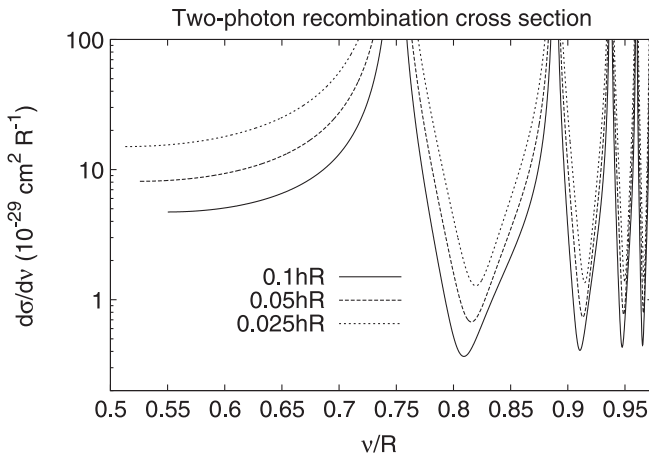


FIG. 2. The differential cross section $d\sigma/d\nu$ for two-photon recombination at three incident electron energies (0.1, 0.05, and $0.025h\mathcal{R}$).

the $n = 3$ levels this gives a two-photon decay rate of 0.0664 times the DG05 estimate, a result of the $4p$, $5p$, etc. intermediate states partially cancelling the $3p$ contribution to \mathcal{M} . WS07 did not have calculated rate coefficients for $n > 3$, so they rescaled all of the DG05 estimates by a factor of 0.0664 and included them in their MLA code. The WS07 method is an improvement, because it partially takes into account the destructive interference in the matrix element. Nevertheless, it still does not take into account all intermediate states and retains the incorrect $d\Lambda/d\nu \propto n$ scaling.

CS08 presented the full calculation of the two-photon decay rate $d\Lambda/d\nu$ including all intermediate states. However, by including these intermediate states they reintroduce the double-counting and reabsorption problems. CS08 argued that the double-counting problem should be solved by defining a “ $1 + 1$ decay profile” ($\phi_{n;s/d}^{1+1\gamma}$ in their notation) consisting of the sum of Lorentzians at each resonance. The decay rate from the $1 + 1$ profile is $\int \phi_{n;s/d}^{1+1\gamma} d\nu = \sum_{n'=2}^{n-1} A_{nl,n'p}$. They then subtract this from the full two-photon decay profile to give a “pure” two-photon profile. Their effective two-photon rate is obtained by integrating this pure two-photon profile

$$\Lambda_{nl}^{\text{CS08}} = \int_{(1-n^{-2})\mathcal{R}/2}^{\nu_c} \left(\frac{d\Lambda_{nl}}{d\nu} - \phi_{n;s/d}^{1+1\gamma} \right) d\nu, \quad (32)$$

where ν_c is a cutoff frequency. (CS08 denote this rate by $A_{n;s/d \rightarrow 1s}^{2\gamma}$.) The “obvious” choice for the cutoff frequency would be the maximum frequency $(1 - n^{-2})\mathcal{R}$, but CS08 argued that a lower cutoff should be used to solve the reabsorption problem. All photons emitted blueward of $\text{Ly}\alpha$ will be reabsorbed later in recombination. Furthermore, photons emitted redward of $\text{Ly}\alpha$ will be reabsorbed if they are within the optically thick part of the red $\text{Ly}\alpha$ damping wing, so ν_c should be somewhat below $\nu_{\text{Ly}\alpha}$. CS08 parameterized this in terms of the number of Doppler widths $x_{\Gamma,c}$

$$\nu_c = \nu_{\text{Ly}\alpha} + x_{\Gamma,c} \sqrt{\frac{2T_r}{m_H c^2}}. \quad (33)$$

(Note the sign convention that $x_{\Gamma,c} < 0$ corresponds to the red side of $\text{Ly}\alpha$.) The pure two-photon rate $\Lambda_{nl}^{\text{CS08}}$ depends only logarithmically on $x_{\Gamma,c}$, because the $1 + 1$ profile removes the order $(\nu - \nu_{\text{Ly}\alpha})^{-2}$ term from the series expansion of the integrand in Eq. (32), so that at small $\nu - \nu_{\text{Ly}\alpha}$ the leading term is of order $(\nu - \nu_{\text{Ly}\alpha})^{-1}$, whose integral is logarithmically divergent. CS08 considered several values of $x_{\Gamma,c}$ ranging from -10^4 to -10^5 , and found that their correction to the free electron abundance was $< 0.5\%$ in each case. They also found that while the full two-photon profile $d\Lambda_{nl}/d\nu$ must be nonnegative, there is no restriction on its sign after the $1 + 1$ profile has been subtracted; their rate coefficient $\Lambda_{nl}^{\text{CS08}}$ is positive for the d initial states and negative for s initial states.

While the CS08 treatment is the most complete thus far, there are three areas in which it could be improved. One is that it contains a free parameter $x_{\Gamma,c}$ that affects the results, and it is not obvious which value should be used. A second issue is the subtraction of the Lorentzian profile for the $1 + 1$ decays. While it is true that integrating the Lorentzian profile does give the one-photon formula $\sum_{n'=2}^{n-1} A_{nl,n'p}$ for the decay rate, the Lorentzian profile plays no role at all in the conventional MLA. The derivation of the Sobolev approximation assumes that the line in question (here $\text{Ly}\alpha$) is a δ function, and hence it is valid for *any* line profile so long as it is sufficiently narrow. Thus, it is not clear mathematically why in Eq. (32) we should use the Lorentzian as the “already included” profile instead of any other function that integrates to $\sum_{n'=2}^{n-1} A_{nl,n'p}$. A final distinct issue is that photons emitted blueward of $\text{Ly}\alpha$ (i.e. in super- $\text{Ly}\alpha$ decays) will be reabsorbed, but *not instantaneously*, so they do have an opportunity to affect recombination; we will see later that they yield corrections of up to $\sim 1.5\%$.

C. Physical resolution of the double-counting and reabsorption problems

To resolve the double-counting and reabsorption problems, in the next section we will proceed to write down the two-photon radiative transfer equations. A radiative transfer calculation is necessary to fully address reabsorption. Both numerical solution of the two-photon equations (Sec. IV) and analytic approximations (Sec. V) will show that the two-photon decay rate and radiation spectrum both remain finite; the photon phase-space density f_ν in the optically thick region near the $\text{Ly}\alpha$ line approaches a quasi-equilibrium value in which the $nl \rightarrow 1s$ decays and $1s \rightarrow nl$ absorptions roughly cancel.

The double-counting problem is conceptually trickier. There is no physical distinction between a pure two-photon decay that happens to be on resonance and a $1 + 1$ decay, so we are free to draw the distinction; however, we want to so long as the total rate and the radiative transfer calculation are correct. Since the MLA is already treating the $1 + 1$ decays and associated radiative transfer via the Sobolev approximation for the Lyman lines, it makes sense for us to set up the distinction in such a way as to preserve the validity of the Sobolev approximation. The assumptions underlying the Sobolev approximation can be identified from e.g. the derivation in Refs. [8,43]; they are (i) that the line width is narrow enough that phase-space factors, atomic density, and Hubble rate do not vary significantly during the passage of a photon through the line; (ii) the emission and absorption profiles are the same; and (iii) that the line is the only significant emission or absorption mechanism in its frequency range. [Additionally the $\text{Ly}\alpha$ diffusion is neglected, but for very optically thick lines the photons in the line tend to come to quasi-equilibrium in which the phase-space density is constant,

$f \approx x_u/(g_u x_l/g_l - x_u)$. In this case, diffusion in frequency plays little role, and the Sobolev escape probability works well. This assumption will be revisited in future work where $\text{Ly}\alpha$ diffusion is included].

Of these, (i) requires that the “1 + 1” decay profile to have support only within a frequency width $\Delta\nu \ll \nu_{\text{Ly}\alpha}$. Assumption (ii) is harder: the ratio of emission to absorption profile in a thermal medium is $\propto e^{-h\nu/T}$ if $h\nu \gg T$. Of course the real Universe is not in thermal equilibrium; however, the emission/absorption ratio exhibits this exponential dependence in the vicinity of $\text{Ly}\alpha$ so long as the excited levels $n \geq 2$ are in Boltzmann equilibrium with each other, even if the Saha equation is not obeyed, and the ratio x_{2p}/x_{1s} is not equal to the Boltzmann value. [We will prove this explicitly in Sec. VB when we derive Eq. (85); the ratio of emission/absorption profiles is f_ν^0]. In order for $e^{-h\nu/T}$ to vary by only a small fraction and satisfy assumption (ii), we need $\Delta\nu \ll T/h$. This “ \ll ” is especially stringent: the Sobolev escape probability is guaranteed to be valid to accuracy ϵ only so long as the emission/absorption ratio ($\propto e^{-h\nu/T}$) is constant to within ϵ , so we need $\Delta\nu < \epsilon T/h$. This condition may be relaxed if the two-photon absorption optical depth (dominated by $1s \rightarrow 3d$) redward of ν_1 (the red boundary of the $\text{Ly}\alpha$ line) is $\gg 1$, since in this case all $\text{Ly}\alpha$ photons that escape the line are reabsorbed anyway, and the escape probability is moot. This issue will be explored numerically when we vary ν_1 .

Assumption (iii) requires that 1 + 1 and pure two-photon decays not overlap in frequency—that is, for each Lyman line, we must define a frequency range $\nu_1 < \nu < \nu_2$. If a two-photon decay produces a photon within this range, it gets counted as 1 + 1. Otherwise (i.e. if the photon is far enough from resonance), it gets counted as a pure two-photon decay. In this picture the “1 + 1” part of the profile resembles a Lorentzian truncated at the frequencies ν_1 and ν_2 , and the far damping wings are considered pure two-photon decays, even if they resemble a Lorentzian shape.

As a practical matter there is another restriction on ν_1 and ν_2 . In our numerical code we are not including the pole displacement in the calculation of the two-photon transition rates, since this would introduce highly complicated temperature dependence into $d\Lambda_{nl}/d\nu$. Therefore, we must choose the detunings $\nu_{\text{Ly}\alpha} - \nu_1$ and $\nu_2 - \nu_{\text{Ly}\alpha}$ to be large compared with the intrinsic width Γ_{2p} of the $\text{Ly}\alpha$ line. This has the ancillary advantage that the 1 + 1 decay rate is not significantly reduced by the truncation of the damping wings: the fraction of decays that proceed through the far damping tails is

$$f_{\text{far}} = 1 - \int_{\nu_1}^{\nu_2} \frac{\Gamma_{2p} d\nu}{4\pi^2[(\nu - \nu_{\text{Ly}\alpha})^2 + (\Gamma_{2p}/4\pi)^2]} \approx \frac{\Gamma_{2p}}{4\pi^2} \left(\frac{1}{\nu_{\text{Ly}\alpha} - \nu_1} + \frac{1}{\nu_2 - \nu_{\text{Ly}\alpha}} \right) \quad (34)$$

in the Lorentzian approximation. This quantity is $\ll 1$ so long as $\nu_{\text{Ly}\alpha} - \nu_1$ and $\nu_2 - \nu_{\text{Ly}\alpha}$ are $\gg \Gamma_{2p}$.

In principle, the $\text{Ly}\alpha$ line profile is affected at some level by two-photon corrections even within the range $\nu_1 < \nu_{\text{Ly}\alpha} < \nu_2$, because the decay process that emits the line (e.g. $3d \rightarrow 2p \rightarrow 1s$) interferes with other pathways (e.g. $3d \rightarrow 3p \rightarrow 1s$) that disrupt the approximation that there is a single pole in the matrix element (which leads to a Lorentzian). This is not a problem for us: as long as assumptions (i)–(iii) above hold, we may use the Sobolev approximation for the central part of the $\text{Ly}\alpha$ line, which does not depend on the line profile.

Note that our separation into 1 + 1 and pure two-photon decays differs from CS08. Also, because ν_1 and ν_2 are arbitrary, one cannot uniquely define a total two-photon decay coefficient Λ_{nl} for the two-photon decays. This is not a problem, since the “total” Λ_{nl} will not appear in our equations, and the notion of this object plays no role in our solution.

Our prescription, like that of CS08, is not unique: ν_1 and ν_2 are arbitrary. The solution we derive can only be correct (and useful) if, within the restrictions above, we get the same answer regardless of our choice of these numbers. This will be verified numerically, and we will also find that the analytic approximations we derive have well-defined limits for small $\nu_{\text{Ly}\alpha} - \nu_1$ and $\nu_2 - \nu_{\text{Ly}\alpha}$ that agree with the numerical results.

D. Two-photon radiative transfer equations

In the last two sections, we highlighted the reasons why it is difficult to graft two-photon transitions onto the MLA. We will now write down the full set of two-photon equations, including those for the radiation field. This full set of equations must include both two-photon emission and absorption in order to be consistent with the laws of thermodynamics and should serve as a starting point for any numerical solution or analytic approximation. After writing down the equations, we will consider two ways of implementing them in an MLA.

In the case of a two-photon decay from $nl \rightarrow 1s$, the net number of decays per unit photon frequency per unit time per unit H atom is

$$\Delta_{nl}(\nu) = \frac{d\Lambda_{nl}}{d\nu} \left[(1 + f_\nu)(1 + f_{\nu'})x_{nl} - \frac{g_{nl}}{g_{1s}} f_\nu f_{\nu'} x_{1s} \right], \quad (35)$$

so that the net two-photon decay rate is

$$\dot{x}_{nl}|_{2\gamma} = - \int_{(1-n^2)\mathcal{R}/2}^{(1-n^2)\mathcal{R}} \Delta_{nl}(\nu) d\nu \quad (36)$$

for $n \geq 2$, and

$$\dot{x}_{1s}|_{2\gamma} = - \sum_{nl, n \geq 2} \dot{x}_{nl}|_{2\gamma}. \quad (37)$$

In the case of the $2s \rightarrow 1s$ transition, these equations replace Eq. (11). Note that it is only necessary to follow the higher-energy photon ($\nu > \nu'$): the lower-energy photon has $\nu' < \mathcal{R}/2$, and hence has negligible probability of re-exciting a hydrogen atom. We can calculate the probability of these photons interacting again by considering their optical depths for various processes. The optical depth per Hubble time for two-photon absorption is typically in the range 10^{-20} – 10^{-16} at frequency $\nu_{\text{Ly}\alpha}/2 = 3\mathcal{R}/8$. The optical depth per Hubble time for Balmer continuum absorption is higher, reaching a peak of 2×10^{-4} at the threshold $\mathcal{R}/4$ and $z = 1400$. The optical depth in the $\text{H}\alpha$ line is 5.5×10^{-4} at $z = 1400$, and since this is a resonance in the Raman-scattering cross section, the optical depth for Raman scattering is much less (typically by a factor of the resonance width divided by the detuning from resonance). Since there are ~ 0.5 net two-photon decays per hydrogen atom during H recombination, the latter numbers suggest that reabsorption of the lower-energy photon from the two-photon decays would influence recombination at the level of a few parts in 10^4 . In practice, the effect should be less since at $z \sim 1400$ the excited states $\text{H}(n \geq 2)$ and the continuum $\text{H}^+ + e^-$ are kept in equilibrium (in which case absorption or emission of Balmer photons has no effect) and at lower redshift the lower population of the $2s$ and $2p$ levels results in much smaller optical depths. Therefore, we have not followed the lower-energy (ν') photons.

Because some portions of the two-photon spectrum are optically thick, we will also need the equations for the photons. For frequencies not on an H I resonance line, the photons follow the equation

$$\dot{f}_\nu = H\nu \frac{\partial f_\nu}{\partial \nu} + \frac{c^3 n_{\text{H}}}{8\pi\nu^2} \sum_{nl} \Delta_{nl}(\nu), \quad (38)$$

where the first term is the Hubble redshift, and the second term is the source (the coefficient describes the photon phase-space density).

One can write similar equations for two-photon recombination/ionization and Raman scattering. Only those reactions involving the ground state are included, because one can transition among the other states by optically thin one-photon transitions, which are much faster. In the case of the Raman-scattering reaction

$$\text{H}(nl) + h\nu' \rightarrow \text{H}(1s) + h\nu, \quad (39)$$

we extend Eq. (35) by writing

$$\Delta_{nl}(\nu) = \frac{dK_{nl}}{d\nu} \left[(1 + f_\nu) f_{\nu'} x_{nl} - \frac{g_{nl}}{g_{1s}} f_\nu (1 + f_{\nu'}) x_{1s} \right] \quad (40)$$

for $\nu > (1 - n^{-2})\mathcal{R}$ and $\nu' = \nu - (1 - n^{-2})\mathcal{R}$. Here, $dK_{nl}/d\nu$ is the differential Raman-scattering rate, defined by Eq. (B10). Equation (36) is extended to include Raman

scattering by writing

$$\dot{x}_{nl}|_{\text{Raman}} = \int_{(1-n^{-2})\mathcal{R}}^{\mathcal{R}} \Delta_{nl}(\nu) d\nu; \quad (41)$$

we have only extended the integration up to \mathcal{R} , because if the output photon has a frequency above \mathcal{R} then it immediately ionizes an H atom. For two-photon recombination to the ground state, i.e.

$$\text{H}^+ + e^- \rightarrow \text{H}(1s) + h\nu + h\nu', \quad (42)$$

one writes

$$\Delta_c(\nu) = \int_0^\infty \left[(1 + f_\nu)(1 + f_{\nu'}) n_{\text{H}} x_e x_p P_{\text{M}}(E_e) - \frac{dg}{dEdV} f_\nu f_{\nu'} \frac{x_{1s}}{2} \right] \alpha^{(2)}(\nu, \nu') dE_e, \quad (43)$$

where $h\nu' = E_e - E_{1s} - h\nu$, and $\alpha^{(2)}(\nu, \nu')$ is the differential two-photon recombination rate coefficient (in e.g. $\text{cm}^3 \text{s}^{-1} \text{Hz}^{-1}$), given by Eq. (B15). The $\Delta_c(\nu)$ term is added along with the bound-bound two-photon transitions $\Delta_{nl}(\nu)$ in Eq. (38).

IV. NUMERICAL APPROACH

In Sec. IIID, we wrote down the radiative transfer and level population equations including two-photon transitions. Now, we solve them numerically by discretizing the two-photon continuum. We first write down the methods of solution (Sec. IVA) and describe our fiducial choice of parameters (Sec. IVB). Then, we consider the results obtained by this method in Sec. IVC, turning on a sequence of effects in order: (A) the “correct” handling of the $2s \rightarrow 1s$ decay, including stimulated emission and feedback; (B) sub-Ly α two-photon decays, i.e. those from $n \geq 3$ where both photon frequencies are below Ly α ; (C) super-Ly α two-photon decays, i.e. those where one-photon frequency is above Ly α ; (D) Raman scattering; and (E) two-photon recombination/ionization.

A. Method

Our agenda here is to discretize the two-photon transfer equations. We will do this by converting the continuous frequency distributions such as $d\Lambda_{nl}/d\nu$ into discrete distributions, i.e. lines. We will define a set of virtual levels, such that a two-photon decay such as $3d \rightarrow 1s$ can be represented formally as a one-photon decay from $3d$ to a virtual level, and then from the virtual level to $1s$. We will then write down a set of equations for these levels so that the equations for the occupation of the virtual level are formally equivalent to the two-photon radiative transfer equations. We emphasize that this is *only* a formal correspondence used to force an MLA code to solve the two-photon transfer equations, and that the virtual levels are *not* actual levels of the hydrogen atom (they are not linearly independent states in the Hilbert space). In order to be

pedagogical, we will work through only the two-photon decay/excitation here, and then state the analogous results for Raman scattering and two-photon recombination/ionization.

In order for the discretization procedure to work, the total rate and (after appropriate smoothing) shape of the two-photon distribution must remain fixed. Since two-photon absorption also plays a key role, we must also make sure that the two-photon optical depths and branching probabilities (relative probability of two-photon absorption to different excited levels) are preserved.

The simplest way of doing this is to multiply each probability distribution by a discretizing function D'

$$\left. \frac{d\Lambda_{nl}}{d\nu} \right|_{\text{used}} = \left. \frac{d\Lambda_{nl}}{d\nu} \right|_{\text{true}} D'(\nu), \quad (44)$$

with

$$D'(\nu) = \sum_{b=1}^{N_\nu} \Delta\nu_b \delta(\nu - \nu_b), \quad (45)$$

with the “spike” frequencies ordered $\nu_1 < \nu_2 < \dots < \nu_{N_\nu}$ and with $\Delta\nu_b$ being the spacing between consecutive spikes. One can also define the integral

$$D(\nu) = \int_{\nu_{\text{cut}}}^{\nu} D'(\tilde{\nu}) d\tilde{\nu} \approx \nu - \nu_{\text{cut}}. \quad (46)$$

As always, the delta function has infinitesimal width. The frequency ν_{cut} is the minimum frequency at which we follow the spectral distortion. Note that $d\Lambda_{nl}/d\nu$ is only defined above half the maximum frequency $\nu > (1 - n^{-2})\mathcal{R}/2$; below this cutoff we will formally take $d\Lambda_{nl}/d\nu = 0$. We take $f_{\nu'}$ to be a blackbody here, i.e. we neglect the spectral distortion at $\nu' < (1 - n^{-2})\mathcal{R}/2$. As shown in Sec. III D, this does not lead to any significant errors.

Now that we have a discretized spectrum, we proceed to solve the radiative transfer equations. At frequencies in between spikes, Eq. (38) is simply a free-streaming equation and has a solution analogous to Eq. (13)

$$f(\nu_b + \epsilon, z) = f\left[\nu_{b+1} - \epsilon, \frac{\nu_{b+1}}{\nu_b}(1 + z) - 1\right]. \quad (47)$$

Within a spike the situation is more subtle. At $\nu \approx \nu_b$, Eq. (38) becomes

$$\begin{aligned} \dot{f}_\nu = H\nu_b \frac{\partial f_\nu}{\partial \nu} + \frac{c^3 n_{\text{H}}}{8\pi\nu_b^3} D'(\nu) \left\{ \sum_{nl} (1 + f_{\nu'}) x_{nl} \frac{d\Lambda_{nl}}{d\nu} \right. \\ \left. - \sum_{nl} \left[\frac{g_{nl}}{g_{1s}} f_{\nu'} x_{1s} - (1 + f_{\nu'}) x_{nl} \right] \frac{d\Lambda_{nl}}{d\nu} f_\nu \right\}. \end{aligned} \quad (48)$$

Since the low-frequency photon comes from the CMB (i.e. we are neglecting two-photon absorption where the low-energy photon is part of the spectral distortion), we may take $f_{\nu'}$ to be constant across the spike (but of course it

does depend on n). Then we may define the coefficients

$$B_1 = \sum_{nl} \left[\frac{g_{nl}}{g_{1s}} f_{\nu'} x_{1s} - (1 + f_{\nu'}) x_{nl} \right] \frac{d\Lambda_{nl}}{d\nu} \quad (49)$$

and

$$B_0 = \sum_{nl} (1 + f_{\nu'}) x_{nl} \frac{d\Lambda_{nl}}{d\nu}. \quad (50)$$

Then Eq. (48) can be written as

$$\frac{\dot{f}_\nu}{H\nu_b} = -\frac{\partial f_\nu}{\partial \nu} + \frac{c^3 n_{\text{H}}}{8\pi H\nu_b^3} D'(\nu) (B_0 - B_1 f_\nu). \quad (51)$$

This equation has an integral solution,

$$\begin{aligned} f_\nu = f_{\nu_b + \epsilon} + \left(\frac{B_0}{B_1} - f_{\nu_b + \epsilon} \right) \\ \times \left(1 - \exp \left\{ \frac{c^3 n_{\text{H}}}{8\pi H\nu_b^3} B_1 [D(\nu) - D(\nu_b + \epsilon)] \right\} \right) \\ + \int_{\nu_b + \epsilon}^{\nu} \frac{\dot{f}_{\tilde{\nu}}}{H\nu_b} \exp \left\{ \frac{c^3 n_{\text{H}}}{8\pi H\nu_b^3} B_1 [D(\nu) - D(\tilde{\nu})] \right\} d\tilde{\nu}. \end{aligned} \quad (52)$$

As we make the delta functions infinitesimally narrow, the integrand in the last term remains finite, but the range of integrated variable becomes infinitesimal. Therefore, this term disappears. We can simplify Eq. (52) by defining

$$\tau_b(\nu) = -\frac{c^3 n_{\text{H}}}{8\pi H\nu_b^3} B_1 [D(\nu) - D(\nu_b + \epsilon)], \quad (53)$$

and

$$\Delta\tau_b \equiv \tau_b(\nu_b - \epsilon) = \frac{c^3 n_{\text{H}}}{8\pi H\nu_b^3} B_1 \Delta\nu_b. \quad (54)$$

Comparing with Eq. (49), this looks similar to a Sobolev optical depth; indeed, $\Delta\nu_b$ is the optical depth to two-photon absorption.

At the end of the day, the relevant numbers to know are the net two-photon decay rate from each excited level via bin b , $\dot{x}_{nl}|_{2\gamma,b}$, and the photon phase-space density at $\nu_b - \epsilon$. The latter can be computed from Eq. (52); it is given by

$$f_{\nu_b - \epsilon} = f_{\nu_b + \epsilon} + \left(\frac{B_0}{B_1} - f_{\nu_b + \epsilon} \right) (1 - e^{-\Delta\tau_b}). \quad (55)$$

The two-photon decay rate can be written as [see Eq. (35)]

$$\begin{aligned} \dot{x}_{nl}|_{2\gamma,b} = - \int_{\nu_b - \epsilon}^{\nu_b + \epsilon} \left[(1 + f_\nu)(1 + f_{\nu'}) x_{nl} - \frac{g_{nl}}{g_{1s}} f_\nu f_{\nu'} x_{1s} \right] \\ \times \frac{d\Lambda_{nl}}{d\nu} D'(\nu) d\nu. \end{aligned} \quad (56)$$

Using Eq. (52) for the phase-space density within the spike, and recalling that the last term is zero, we see that

$$\dot{x}_{nl}|_{2\gamma,b} = -\frac{d\Lambda_{nl}}{d\nu}[(1 + \bar{f}_{\nu_b})(1 + f_{\nu'})x_{nl} - \bar{f}_{\nu_b}f_{\nu'}x_{1s}]\Delta\nu_b, \quad (57)$$

where the weighted mean phase-space density has been defined as

$$\bar{f}_{\nu_b} = \frac{B_0}{B_1} + \left(f_{\nu_b+\epsilon} - \frac{B_0}{B_1}\right) \frac{1 - e^{-\Delta\tau_b}}{\Delta\tau_b}. \quad (58)$$

It is convenient to define the Sobolev-like “escape probability” from the spike

$$\Pi_b \equiv \frac{1 - e^{-\Delta\tau_b}}{\Delta\tau_b}. \quad (59)$$

In principle, we could directly add Eqs. (55) and (57) to the level equations, since \bar{f}_{ν_b} is linear in the excited state populations so long as they are small enough that x_{1s} dominates the sum in B_1 [Eq. (49)]. We want to do this in a way that is easily incorporated into the steady-state equations; this can be done by inventing a “virtual” level b for each spike, and extending the steady-state source vector \mathbf{s} and transition matrix \mathbf{T} to include it in such a way as to reproduce the two-photon equations. The populations of the virtual level for each spike will correspond directly to the spectral distortion, which will be convenient numerically. Let us define for each spike an arbitrary “degeneracy” g_b . Formally, one should think of g_b as infinitesimal, since the virtual levels cannot be occupied, but within the context of the steady-state approximation the argument below will work for any arbitrary value. We will then define the virtual level occupation

$$x_b \equiv \frac{g_b}{2} x_{1s} \bar{f}_{\nu_b} \quad (60)$$

and the virtual level energy

$$E_b \equiv E_{1s} + h\nu. \quad (61)$$

We will also define the virtual rate coefficient

$$A_{nl,b} = \frac{d\Lambda_{nl}}{d\nu} \Big|_{\nu=\nu_b} \Delta\nu_b. \quad (62)$$

Then Eq. (57) gives us the effect of a given spike on the rate equation for level nl . Since $f_{\nu'} = f_{nl,b+}$ and $\bar{f}_{\nu_b} \ll 1$, this is

$$\dot{x}_{nl}|_b = -A_{nl,b}(1 + f_{nl,b+})x_{nl} + \frac{g_{nl}}{g_b} A_{nl,b} f_{nl,b+} x_b. \quad (63)$$

This is exactly equal to what Eq. (18) for R_{ji} would give when plugged into the steady-state equation, Eq. (16).

In order to implement something like the steady-state equations, we must also express the equation for \bar{f}_{ν_b} [Eq. (58)] in language that allows us to define a \mathbf{T} matrix.

That is, if we imagine extending Eq. (21) to include both real (r) and virtual (v) levels

$$\mathbf{0} = \begin{pmatrix} \mathbf{s}^r \\ \mathbf{s}^v \end{pmatrix} - \begin{pmatrix} \mathbf{T}^{rr} & \mathbf{T}^{rv} \\ \mathbf{T}^{vr} & \mathbf{T}^{vv} \end{pmatrix} \begin{pmatrix} \mathbf{x}^r \\ \mathbf{x}^v \end{pmatrix}, \quad (64)$$

we want to construct values of \mathbf{s}^v , \mathbf{T}^{vr} , and \mathbf{T}^{vv} that are equivalent to Eq. (58). (Note that we have already constructed the top row of the matrices.) Rewriting Eq. (58) in the form

$$0 = -\left(\sum_{nl} f_{\nu'} \frac{g_{nl}}{g_b} A_{nl,b}\right) x_b + (1 - \Pi_b) \sum_{nl} (1 + f_{\nu'}) A_{nl,b} x_{nl} + \left(\sum_{nl} f_{\nu'} \frac{g_{nl}}{g_b} A_{nl,b}\right) \frac{g_b}{g_{1s}} x_{1s} f_{\nu_b+\epsilon} \Pi_b, \quad (65)$$

we can then define the virtual level source term

$$s_b = \left(\sum_{nl} f_{\nu'} \frac{g_{nl}}{g_b} A_{nl,b}\right) \frac{g_b}{g_{1s}} x_{1s} f_{\nu_b+\epsilon} \frac{\Pi_b}{1 - \Pi_b}, \quad (66)$$

the virtual-real sub-block of the \mathbf{T} -matrix,

$$T_{b,nl} = -\sum_{nl} (1 + f_{\nu'}) A_{nl,b}, \quad (67)$$

and the virtual-virtual sub-block,

$$T_{bb'} = \frac{\delta_{bb'}}{1 - \Pi_b} \sum_{nl} f_{\nu'} \frac{g_{nl}}{g_b} A_{nl,b}. \quad (68)$$

A simple algebraic calculation shows that this is precisely what one would calculate for s_b , $T_{b,nl}$, and $T_{bb'}$ if one were to make the replacement

$$A_{b,1s} P_{b,1s} \rightarrow \frac{\Pi_b}{1 - \Pi_b} \sum_{nl} \frac{g_{nl}}{g_b} f_{\nu'} A_{nl,b} \quad (69)$$

in Eqs. (17) and (20). This can also be rewritten as

$$A_{b,1s} P_{b,1s} \rightarrow \frac{8\pi H \nu_b^3}{c^3 n_H} \frac{1 - e^{-\Delta\tau_b}}{1 - (1 - e^{-\Delta\tau_b})/\Delta\tau_b}. \quad (70)$$

In other words, the steady-state equations can be extended to include the two-photon decays by adding a set of virtual levels, inserting a virtual one-photon transition between each level nl and a virtual level with Einstein coefficient given by the two-photon decay rate, and introducing a transition from the virtual level to the ground state $1s$ with rate given by Eq. (70).

The above methodology can be extended to include Raman scattering and two-photon recombination/ionization with no conceptual changes but with the addition of some new reaction rates. The Raman-scattering reaction, Eq. (B7), is included by including a new radiative rate

$$A_{b,nl} = \frac{g_{nl}}{g_b} \frac{dK_{nl}}{d\nu} \Delta\nu_b, \quad (71)$$

and two-photon recombination is included by writing an effective recombination coefficient to the virtual level

$$\alpha_b(E_e) = \alpha^{(2)}(\nu, \nu') \Delta \nu_b, \quad (72)$$

where $h\nu' = E_e - E_{1s} - h\nu$. Equation (70) remains valid as long as we define a total optical depth

$$\begin{aligned} \Delta \tau_b = & \frac{c^3 n_H}{8\pi H \nu_b^3} x_{1s} \left[\sum_{nl > b} \frac{g_{nl}}{g_{1s}} f_{\nu'} \frac{d\Lambda_{nl}}{d\nu} \right. \\ & + \sum_{nl < b} \frac{g_{nl}}{g_{1s}} (1 + f_{\nu'}) \frac{dK_{nl}}{d\nu} + \frac{2^{7/2} \pi \mu^{3/2}}{h^3} \\ & \left. \times \int_0^\infty dE_e \alpha^{(2)}(\nu, \nu') \frac{E_e}{g_{1s}} f_{\nu'} dE_e \right]. \quad (73) \end{aligned}$$

B. Parameters

We have run our numerical radiative transfer code several times with different sets of parameters. These choices are summarized in Table I. The truncation frequencies of the Ly α line (outside of which two-photon decays are treated using the numerical radiative transfer instead of being treated as 1 + 1 decays with the Sobolev approximation) are $\nu_{1,2} = \nu_{\text{Ly}\alpha} \pm \Delta \nu_{\text{max}}$. In the latter case, one should compare with the intrinsic width of the $2p$ intermediate state, $\Gamma_{2p} = A_{2p,1s} = 6.3 \times 10^8 \text{ s}^{-1}$ in vacuum. (The additional finite-temperature broadening is due mainly to $2p \rightarrow 3s, 3d$ transitions and for recombination-era temperatures is $\ll A_{2p,1s}$.) Thermal equilibrium conditions were assumed for the initial data at z_{init} . This redshift was chosen to be after the completion of He II \rightarrow He I recombination in order to avoid needing to simultaneously follow hydrogen and helium. The “basic” set of parameters will be used for all plots in the paper unless otherwise indicated. All runs commenced at $z_{\text{init}} = 1605.8$ and used $n_{\text{max}} = 30$ shells of the hydrogen atom (465 sub-levels including l resolution).

For the Basic run, the frequency grid extends only up to $\text{Ly}\iota = 0.99\mathcal{R}$, since above this frequency the real hydrogen lines are very closely spaced and feedback is almost instantaneous. The “WideLine 1/2” runs are the same as the Basic run except that we increased $\Delta \nu_{\text{max}}$ by a factor of 3 and 9, respectively, i.e. a two-photon decay had to emit a photon within 315 (instead of 105) GHz of the Ly α line in order to be considered a 1 + 1 decay instead of being treated with the two-photon radiative transfer. Its purpose is to show that $\Delta \nu_{\text{max}}$ does not matter. The “HalfStep” run is a convergence test that used half the time step of the Basic run. The “LoRes” run has a lower-resolution grid of virtual levels, while “HiRes” has a higher-resolution grid (twice the resolution except within 25 THz of Ly α) that extends all the way up to Ly μ ($1s - 13p$). The “CoarseBF” run uses a step size of $\Delta \ln E_e = 0.2$ instead of 0.1 for evaluating the energy integrals for the bound-free rates. All of the alternative runs gave changes in the ionization fraction $|\Delta x_e|/x_e < 10^{-4}$, except WideLine 2 and LoRes for which the maximum change was 4×10^{-4} .

C. Results

The change in the recombination history caused by the new physics is shown in Fig. 3. We will discuss the effects qualitatively here; analytic estimates will be given in Sec. V and compared with the numerical results. The strength of the numerical result is that the only approximations made are those intrinsic to the two-photon radiative transfer equation and the step sizes, frequency resolution, etc., which can be varied to establish convergence. The analytic approximations yield insight but require numerous additional simplifying assumptions whose validity is difficult to assess.

The first effect that we have considered is the inclusion of stimulated two-photon decay and nonthermal two-photon absorption in the $2s \rightarrow 1s$ transition (A). These effects go in opposite directions, since stimulated emission increases the effective $2s \rightarrow 1s$ decay rate and hence accelerates recombination [23], whereas nonthermal two-

TABLE I. The parameters for each run of the radiative transfer code. Shown are the number of virtual levels N_{virt} ; and the step size in log scale factor $\Delta \ln a$; the maximum detuning from Ly α at which a photon is considered resonant $\Delta \nu_{\text{max}}$; the step size in the bound-free rate integrals $\Delta \ln E_e$; and the line corresponding to the maximum frequency used in the grid.

Run	N_{virt}	$10^5 \Delta \ln a$	$\Delta \nu_{\text{max}}$ GHz	$\Delta \ln E_e$	Max. grid frequency
Basic	367	4.25	105	0.1	Ly ι
WideLine1	365	4.25	315	0.1	Ly ι
WideLine2	365	4.25	945	0.1	Ly ι
LoRes	262	4.25	6116 ^a	0.1	Ly ι
HalfStep	367	2.125	105	0.1	Ly ι
HiRes	695	4.25	105	0.1	Ly μ
CoarseBF	367	4.25	105	0.2	Ly ι

^aFor the red boundary ν_1 ; the blue boundary is at $\nu_{\text{Ly}\alpha} + 10259 \text{ GHz}$.

photon absorption retards it [24]. Initially, the net effect is a slight acceleration because of the strong stimulated emission at high temperature, but by $z = 1345$ this turns into a retardation ($\Delta x_e > 0$), as enough hydrogen recombines to build a significant Ly α spectral distortion. This grows to $\Delta x_e \approx 0.001$ at $z \approx 1200$ and then tapers off. For most redshifts we are in agreement with Ref. [24] (compare their Fig. 9 to our Fig. 3), but at the tail end of recombination ($z < 800$) we find a smaller effect, $\Delta x_e/x_e = 0.0018$ at $z = 700$ instead of 0.0047. This is probably due to the approximation in Ref. [24] that the excited levels $n \geq 2$ are in Saha equilibrium; if we force Saha equilibrium in our code, our $\Delta x_e/x_e$ rises to 0.0048.

Next, we considered (B), the sub-Ly α two-photon transitions $ns, nd \leftrightarrow 1s$. In these cases, the emitted photons can never be resonantly absorbed by a hydrogen atom. As one would intuitively expect, the addition of this new process brings the Universe closer to thermal equilibrium, that is, it accelerates recombination. These transitions have a huge effect: they reduce the electron abundance by as much as $\sim 2\%$, and moreover this reduction is occurring at $z \sim 1100$, i.e. at the surface of last scattering where it matters most.

The next effect is (C), the super-Ly α two-photon decays. This is another process connecting excited states of H I to the ground state and naively we would expect that like (B) it would accelerate recombination. This is true at first. However, at $z = 1290$ the situation reverses, and the new process delays recombination. The reason for this is that these two-photon decays are populating the region of photon phase-space at $\nu > \nu_{\text{Ly}\alpha}$ to far above the blackbody value. In the later stages of recombination, these photons redshift down to Ly α and begin re-exciting hydrogen atoms. This delayed feedback is present even in the standard recombination calculation, because some photons escape from the Ly β line [40], but the two-photon decays to frequencies between Ly α and Ly β make a significant addition, as their probability of immediate reabsorption is much lower. A similar result occurs for (D) Raman scattering, which initially accelerates recombination but then slows it down. In this case, the main effect is $2s \rightarrow 1s$ scattering in which the incoming photon is at $\nu < \nu_{\text{Ly}\alpha}$ and the outgoing photon is at $\nu_{\text{Ly}\alpha} < \nu < \nu_{\text{Ly}\beta}$.

Two-photon recombinations (E) have an accelerating effect as one would intuitively expect. However, since the electron occupation probability of the continuum states ($E_e > 0$) is much less than the $2s$, $3s$, and $3d$ states, this is a small effect; we find a maximum of $|\Delta x_e|/x_e \approx 10^{-3}$.

We can also examine the effect on the high-frequency end of the spectral distortion. In Fig. 4, we show the photon spectrum at three epochs during recombination. The full calculation including effects A–E is shown compared with the case with effect A (modifications to $2s \leftrightarrow 1s$) only. (We show the case with the $2s \leftrightarrow 1s$ modifications turned on, because in the original case with the $2s \rightarrow 1s$ decays

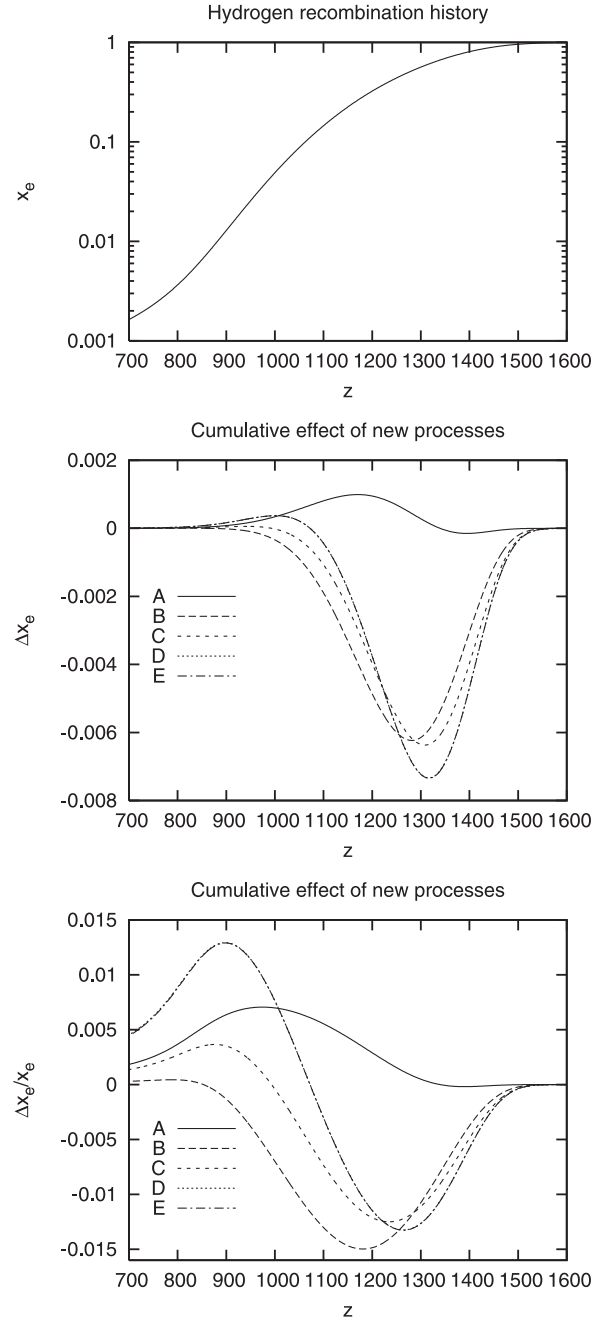


FIG. 3. Upper panel: the recombination history. Middle panel: the absolute change Δx_e caused by incorporation of each process. Lower panel: the relative change $\Delta x_e/x_e$. The changes are cumulative in the sense that the “B” curve represents the effect of incorporating A and B, etc. The (D) and (E) curves are almost indistinguishable on the plot.

followed with the integrated rate coefficient Λ_{2s} only, the code cannot follow the two-photon spectral distortion.) The large jumps at each Lyman line are due to photons redshifting out of the line, as described by Peebles [5]. However, even at early times there is a spectral distortion *between* the Lyman lines that is due to two-photon and

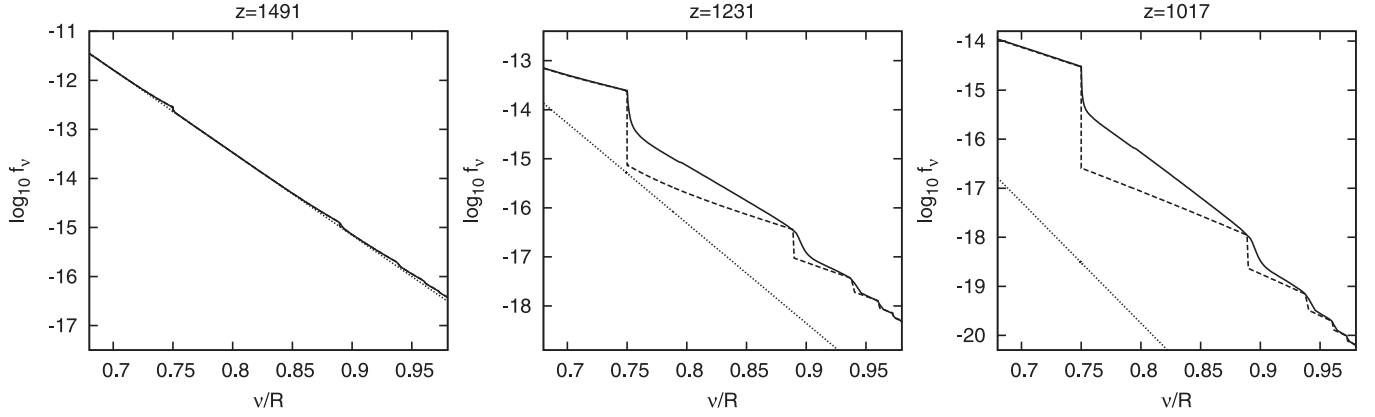


FIG. 4. The radiation spectrum as a function of time. Each panel shows the full spectrum (solid line); the spectrum neglecting the highly excited two-photon and Raman transitions (i.e. including only modification A) (dashed line); and the blackbody spectrum (dotted line). We do not show the lower frequencies, because our code does not follow the lower-energy photon in two-photon decays nor the Balmer, Paschen, etc. lines or bound-free continuum.

Raman transitions. Particularly impressive is the phase space between $\text{Ly}\alpha$ and $\text{Ly}\beta$: here, the photon occupation number can be up to an order of magnitude larger than in the traditional treatment. Because of their very long mean free paths, these photons can survive to $z < 1100$ and slow down recombination by re-exciting hydrogen atoms.

V. ANALYTIC APPROACH

In the previous section, we numerically evaluated the effect of two-photon transitions on hydrogen recombination using a radiative transfer scheme based on virtual levels. Here, we consider simple analytic estimates of their effect. The main purposes are to provide an independent check of the fully numerical result and to develop some understanding of the magnitude of the various effects. We consider the improvements to the $2s \leftrightarrow 1s$ decays (Sec. VA); sub- $\text{Ly}\alpha$ two-photon decays (Sec. VB); super- $\text{Ly}\alpha$ two-photon decays (Sec. VC); and Raman scattering (Sec. VD). We also considered the two-photon recombinations in the numerical code; however, since they were found to be unimportant, we will not construct an analytic estimate.

Our analytic investigation of the two-photon processes is based on many simplifying assumptions, most notably that (1) $T_r \ll h\mathcal{R}$; (2) the low-lying excited states of the hydrogen atom ($n = 2, 3$) are in Boltzmann equilibrium; and (3) that the radiation field can be treated as being in steady state (or some lowest-order non-steady-state approximation, in the cases of two-photon decays, where one photon is above $\text{Ly}\alpha$ and Raman scattering). These assumptions are only marginally satisfied during recombination, and it is difficult to directly estimate the residual error. None of these approximations is made in the numerical code. For these reasons, we believe the numerical results in the previous section are the most reliable estimate of the effect of two-photon transitions.

A. Stimulated emission and nonthermal absorption in $2s \leftrightarrow 1s$

The $2s \leftrightarrow 1s$ transition is modified by both stimulated emission and absorption of the Lyman- α spectral distortion. The approach described here, like that of Refs. [23,24], assumes the optically thin limit for the $1s \rightarrow 2s$ absorption, but also takes the leading-order term in the power series for the spontaneous decay rate in order to arrive at an analytic formula. This effect requires a thermal CMB photon to be present at one of the photon frequencies ν' and hence occurs in the regime $h\nu' \sim T_r \ll h\mathcal{R}$. In this regime the $2p$ intermediate state dominates; using the exact hydrogenic matrix elements $\langle 1s || \mathbf{r} || 2p \rangle = 2^{15/2} a_0 / 3^{9/2}$ and $\langle 2s || \mathbf{r} || 2p \rangle = -3^{3/2} a_0$, we find [see Eq. (B4)]

$$\frac{d\Lambda_{2s}}{d\nu} \approx \frac{512\alpha_{\text{fs}}^6}{729} \frac{\nu'}{\mathcal{R}}. \quad (74)$$

The total rate $\Delta\dot{x}_1^{(A)}$ of stimulated emission minus absorption of the spectral distortion is

$$\Delta\dot{x}_1^{(A)} = \int_0^{\nu_{\text{Ly}\alpha}/2} \frac{d\Lambda}{d\nu} f_{\nu'} (x_{2s} - x_{1s} \Delta f_{\nu}) d\nu', \quad (75)$$

where Δf_{ν} is the phase-space density of spectral distortion photons (i.e. the total phase-space density with the blackbody subtracted). Since $f_{\nu'}$ is a blackbody function, we may write

$$\Delta\dot{x}_1^{(A)} \approx \frac{512\alpha_{\text{fs}}^6}{729\mathcal{R}} \int_0^{\nu_{\text{Ly}\alpha}/2} \frac{\nu' (x_{2s} - x_{1s} \Delta f_{\nu})}{e^{h\nu'/T_r} - 1} d\nu'. \quad (76)$$

Here, x_{2s} and x_{1s} are independent of frequency, and Δf_{ν} varies smoothly with ν' (since we are looking on the red side of $\text{Ly}\alpha$; if we considered negative ν' there would be a jump). Therefore, this integral can be approximated using one-point Gaussian integration (see, e.g. Chapter 19 of

Ref. [44].) The idea of one-point Gaussian integration is that if one has an analytically integrable weight function $w(\nu)$, and a linear function $g(\nu') = a + b\nu'$, that

$$\int w(\nu')g(\nu')d\nu' = C_0g(\nu'_0), \quad (77)$$

where $C_0 = \int w(\nu')d\nu'$ and $\nu'_0 = C_0^{-1} \int \nu' w(\nu')d\nu'$. This enables integration with a single sample of the function g instead of two as is usual for linear functions. For functions $g(\nu')$ that vary smoothly and are being integrated over a narrow range, we may use Eq. (77) as an approximation. In our case, $w(\nu') = \nu'/(e^{h\nu'/T_r} - 1)$ and $g(\nu') = x_{2s} - x_{1s}\Delta f_{\nu'}$. Application of Gaussian integration gives

$$\int_0^{\nu_{Ly\alpha}/2} \frac{\nu'(x_{2s} - x_{1s}\Delta f_{\nu'})d\nu'}{e^{h\nu'/T_r} - 1} \approx C_0(x_{2s} - x_{1s}\Delta f_{\nu'_0}), \quad (78)$$

where

$$C_0 = \int_0^\infty \frac{\nu'd\nu'}{e^{h\nu'/T_r} - 1} = \frac{\pi^2 T_r^2}{6h^2}, \quad (79)$$

and

$$\nu'_0 = C_0^{-1} \int_0^\infty \frac{\nu'^2 d\nu'}{e^{h\nu'/T_r} - 1} = \frac{12\zeta(3)T_r}{\pi^2 h}. \quad (80)$$

(The upper cutoff in the integral does not matter, because in practice the exponential provides a cutoff.) Here, $\zeta(3)$ is the Riemann ζ function, and $\nu_0 \equiv \nu_{Ly\alpha} - \nu'_0$. This leads to

$$\Delta \dot{x}_1^{(A)} \approx \frac{256\pi^2 \alpha_{is}^6 T_r^2}{2187h^2 \mathcal{R}} (x_{2s} - x_{1s}\Delta f_{\nu_0}). \quad (81)$$

We have compared Eq. (81) to the full numerical result for the $2s \leftrightarrow 1s$ two-photon transitions in Fig. 5.

B. Sub-Ly α two-photon decays

We now consider the sub-Ly α two-photon decays from higher excited states of H I such as $3s$ and $3d$. In most

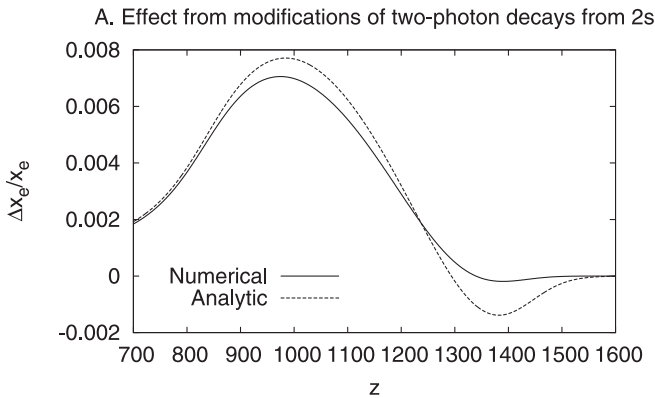


FIG. 5. A comparison of methods for handling the $2s \leftrightarrow 1s$ two-photon transitions. We show the numerical radiative transfer method (solid line) and the analytic approximation of Eq. (81) (dashed line).

cases, the higher-frequency photon is just below Ly α and can be considered to be part of the red damping tail. Therefore, in order to construct an analytic theory of the two-photon transitions we must develop an approximation for the radiative transfer in the red damping tail. The approximation described here draws heavily on that of Appendix B of Hirata and Switzer [22].

The two-photon transfer equation, Eq. (38) with Eq. (35) used for $\Delta_{nl}(\nu)$, can be written as a linear equation for f

$$\frac{\dot{f}_\nu}{H\nu} = \frac{\partial f_\nu}{\partial \nu} - \kappa_\nu(f_\nu - f_\nu^0), \quad (82)$$

where we have defined the functions

$$\kappa_\nu = \frac{c^3 n_H}{8\pi H \nu^3} \sum_{nl} \left[\frac{g_{nl}}{g_{1s}} f_{\nu'} x_{1s} - (1 + f_{\nu'}) x_{nl} \right] \frac{d\Lambda_{nl}}{d\nu} \quad (83)$$

and

$$f_\nu^0 = \frac{\sum_{nl} (1 + f_{\nu'}) x_{nl}}{\sum_{nl} [(g_{nl}/g_{1s}) f_{\nu'} x_{1s} - (1 + f_{\nu'}) x_{nl}]}. \quad (84)$$

Here, κ_ν can be thought of as an optical depth per unit frequency (this is more convenient in cosmology than optical depth per g cm^{-2} as is usual in stellar structure), and f_ν^0 is the photon phase-space density that one would have if the level populations were maintained but the density were taken to ∞ (high optical depth limit). Note that the “optical depth” used here is only the optical depth to two-photon transitions and does not include Rayleigh or Thomson scattering.

Our goal here is to develop an analytic approximation to Eq. (82) valid in the red damping wing of Ly α . To do this we make three assumptions. We begin by making the steady-state approximation that the time derivative on the left-hand side can be dropped. Physically, this corresponds to the statement that the radiation spectrum in the vicinity of Ly α changes slowly compared with the time a photon takes to redshift out of the optically thick part of the damping wing. Secondly, we assume that the excited levels ($n \geq 2$) are in equilibrium with $2p$, i.e. x_{nl}/x_{2p} is given by the Boltzmann factor $[(2l+1)/3]e^{-(E_{nl}-E_{2p})/T_r}$. This, combined with $x_{2p} \ll x_{1s}$, allows us to write

$$f_\nu^0 \approx \frac{x_{2p}}{3x_{1s}} e^{h(\nu - \nu_{Ly\alpha})/T_r}. \quad (85)$$

Our third assumption is that we may approximate $d\Lambda_{nl}/d\nu$ by the single pole approximation $d\Lambda_{nl}/d\nu \propto (\nu - \nu_{Ly\alpha})^{-2}$ in which the two-photon differential decay rate is proportional to the inverse-square of the detuning from the Ly α resonance. Because of the pole in the two-photon matrix element \mathcal{M} from the $2p$ intermediate state, this is the leading-order term in the power series expansion of $d\Lambda_{nl}/d\nu$ around $\nu = \nu_{Ly\alpha}$. The coefficient of this pole is taken from Eq. (B5): in the vicinity of Ly α , we have

$$\frac{d\Lambda_{nl}}{d\nu} \approx \frac{27\alpha_{\text{fs}}^6 \nu_{2p,1s}^3 \nu_{nl,2p}^3}{a_0^4 \mathcal{R}^6 (\nu - \nu_{\text{Ly}\alpha})^2} |\langle nl || \mathbf{r} || 2p \rangle \langle 2p || \mathbf{r} || 1s \rangle|^2. \quad (86)$$

It follows that

$$\kappa_\nu \approx \frac{\bar{W}}{(\nu - \nu_{\text{Ly}\alpha})^2} e^{-h(\nu - \nu_{\text{Ly}\alpha})/T_r}, \quad (87)$$

where

$$\bar{W} \approx \frac{c^3 n_{\text{H}} x_{1s}}{8\pi H \nu_{2p,1s}^3} \sum_{nl, n \geq 3} \frac{2l+1}{e^{h\nu_{nl,2p}/T_r} - 1} \times \frac{27\alpha_{\text{fs}}^6 \nu_{2p,1s}^3 \nu_{nl,2p}^3}{a_0^4 \mathcal{R}^6} |\langle nl || \mathbf{r} || 2p \rangle \langle 2p || \mathbf{r} || 1s \rangle|^2. \quad (88)$$

[We have neglected x_{nl} in Eq. (83), since even with the factor of $1 + f_\nu$ instead of f_ν it will be negligible compared with x_{1s} ; this is physically equivalent to neglecting two-photon decays where the higher-energy photon is stimulated]. This expression can be reconstructed in terms of Einstein coefficients

$$\bar{W} \approx \frac{c^3 n_{\text{H}} A_{2p,1s} x_{1s}}{8\pi H \nu_{2p,1s}^3} \sum_{nl, n \geq 3} \frac{2l+1}{e^{h\nu_{nl,2p}/T_r} - 1} \frac{A_{nl,2p}}{4\pi^2}. \quad (89)$$

The prefactor is recognizable as the Ly α (Gunn-Peterson) optical depth

$$\tau_{\text{Ly}\alpha} \equiv \frac{3c^3 n_{\text{H}} A_{2p,1s} x_{1s}}{8\pi H \nu_{2p,1s}^3}, \quad (90)$$

so we write

$$\bar{W} \approx \frac{\tau_{\text{Ly}\alpha}}{4\pi^2} \sum_{nl, n \geq 3} \frac{2l+1}{3} \frac{A_{nl,2p}}{e^{h\nu_{nl,2p}/T_r} - 1}. \quad (91)$$

With this set of approximations, the radiative transfer equation reduces to

$$\frac{df_\nu}{d\nu} = \frac{\bar{W}}{(\nu - \nu_{\text{Ly}\alpha})^2} \left[e^{h(\nu - \nu_{\text{Ly}\alpha})/T_r} f_\nu - \frac{x_{2p}}{3x_{1s}} \right]. \quad (92)$$

The change of variables $y = h(\nu - \nu_{\text{Ly}\alpha})/T_r$ and $f_\nu = x_{2p}\Phi(y)/(3x_{1s})$ gives

$$\frac{d\Phi}{dy} = \frac{W}{y^2} (e^y \Phi - 1), \quad (93)$$

where $W \equiv h\bar{W}/T_r$ is a dimensionless number that quantifies the normalization of the two-photon opacity. We can understand its physical significance from Eq. (87). For $W \ll 1$ the integrated two-photon optical depth at frequencies less than some critical value $\nu_c < \nu_{\text{Ly}\alpha}$ is

$$\tau_{<} = \int_{\nu_c}^{\nu_{\text{Ly}\alpha}} \kappa_\nu d\nu \approx \frac{T_r}{h} \frac{W}{\nu_{\text{Ly}\alpha} - \nu_c}, \quad (94)$$

so two-photon absorption is optically thick out to a detuning $\sim WT_r/h$ from the Ly α line, and in the optically thick

region the exponential factor in Eq. (92) is a perturbation. The physical situation realized in recombination is always $W < 1$ —we find a maximum of $W = 0.034$ —but the corrections due to nonnegligible W are important for the accuracy desired in upcoming CMB experiments.

Equation (93) forces the boundary condition $\Phi(y = 0) = 1$, and the equation can be evolved by integrating to the left (negative y). It is readily seen that as $W \rightarrow 0$ this boundary condition at $\Phi(y = 0)$ forces $\Phi(-\infty) \rightarrow 1$. The numerical integration of Eq. (93) with an implicit ODE solver is extremely fast compared with the rate coefficient evaluations and linear algebra required when solving recombination, so we have simply solved Eq. (93) each time a value of $\Phi(-\infty)$ is needed. [Since $\Phi(-\infty)$ depends only on the single independent variable W , one could generate a table of values and use spline interpolation if the need arose for a faster evaluation.] The numerical solution is shown in Fig. 6.

The net rate of two-photon transitions to $1s$ that involve photons in the red tail of Ly α can be obtained as follows. We compare Eq. (36), which gives the net $nl \rightarrow 1s$ rate via these transitions, to Eq. (38), which gives the rate of production of photons in the red tail of Ly α . Integrating the latter from frequency ν_1 to $\nu_{\text{Ly}\alpha}$, and making the steady-state approximation again to remove \dot{f}_ν , gives

$$0 = H \int_{\nu_1}^{\nu_{\text{Ly}\alpha}} \nu \frac{\partial f_\nu}{\partial \nu} d\nu + \sum_{nl} \int_{\nu_1}^{\nu_{\text{Ly}\alpha}} \frac{c^3 n_{\text{H}}}{8\pi \nu^2} \Delta_{nl}(\nu) d\nu. \quad (95)$$

Now, if almost all of the two-photon decays from the $n \geq 3$ levels emit one of the photons just below Ly α , we may take $\nu \approx \nu_{\text{Ly}\alpha}$ in the prefactor of the first integral and the denominator of the second integral. Then,

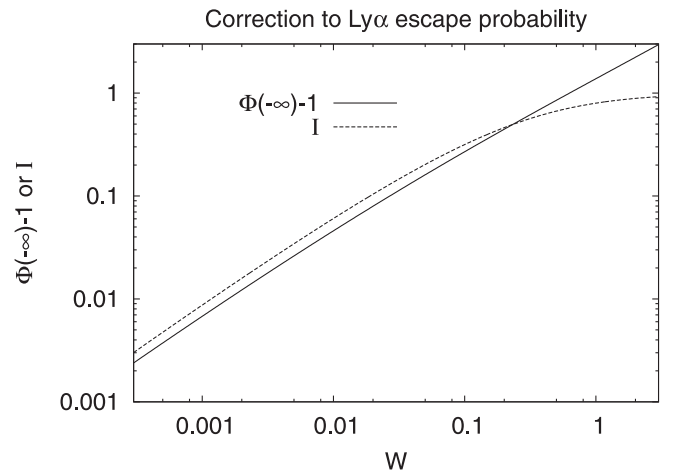


FIG. 6. Solid line: the correction to the Ly α escape probability $\Phi(-\infty)$ as a function of W . Dashed line: the integral I that describes the number of photons on the blue side of Ly α emitted in two-photon decays. Note that as $W \rightarrow 0$ both $\Phi(-\infty) - 1$ and $I \rightarrow 0$.

$$0 = H\nu_{\text{Ly}\alpha}(f_{\nu_{\text{Ly}\alpha}} - f_{\nu_1}) + \frac{c^3 n_{\text{H}}}{8\pi\nu_{\text{Ly}\alpha}^2} \sum_{nl} \int_{\nu_1}^{\nu_{\text{Ly}\alpha}} \Delta_{nl}(\nu) d\nu. \quad (96)$$

The last integral is the net rate of two-photon decays with the higher-energy photon emitted between ν_1 and $\text{Ly}\alpha$. (Here “net” means with the two-photon absorptions subtracted.) Since $\Delta_{nl}(\nu)$ is strongly peaked near $\text{Ly}\alpha$ due to the shape of the two-photon spectrum, we may take ν_1 to be redward of the significant two-photon emission, equate the last integral with $\dot{x}_{nl}|_{2\gamma, \nu < \nu_{\text{Ly}\alpha}}$, and then have

$$0 = H\nu_{\text{Ly}\alpha}(f_{\nu_{\text{Ly}\alpha}} - f_{\nu_1}) + \frac{c^3 n_{\text{H}}}{8\pi\nu_{\text{Ly}\alpha}^2} \Delta\dot{x}_{\downarrow}^{(\text{B})}, \quad (97)$$

where $\Delta\dot{x}_{\downarrow}^{(\text{B})}$ is the net rate of two-photon transitions to $1s$ that involve photons in the red tail of $\text{Ly}\alpha$. Recalling the definition of $\Phi(y)$, that $\nu_{\text{Ly}\alpha}$ and ν_1 correspond to $y = 0$ and $y = -\infty$, respectively, and that $\Phi(0) = 1$, we find

$$\Delta\dot{x}_{\downarrow}^{(\text{B})} = \frac{8\pi H\nu_{\text{Ly}\alpha}^3}{c^3 n_{\text{H}}} \frac{x_{2p}}{3x_{1s}} [\Phi(-\infty) - 1]. \quad (98)$$

This analytical approximation is not a surprise: it is just an extension of the Peebles [5] argument that the net decay rate via emission/absorption of photons in a narrow region of the electromagnetic spectrum that is taken to be in steady state is equal to the difference in phase-space densities on the two sides of the region, multiplied by the appropriate normalization factors. In fact, Eq. (98) without the factor $\Phi(-\infty) - 1$ (and with a correction for black-body photons entering on the blue side) is the Peebles formula for the $\text{Ly}\alpha$ decay rate, so we can think of $\Phi(-\infty)$ as representing an effective enhancement of the $\text{Ly}\alpha$ production rate due to two-photon transitions into the red damping wing. In our approximation, this enhancement factor depends only on W .

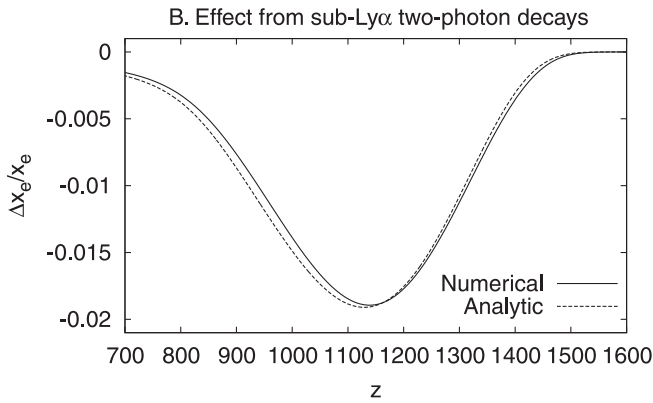


FIG. 7. A comparison of methods for handling the sub- $\text{Ly}\alpha$ $n \geq 3$ two-photon decays. We show the numerical radiative transfer method (solid line) and the analytic approximation of Eq. (99) (dashed line).

We now have all of the pieces required to incorporate the additional decay rate given by Eq. (98) into our recombination model. Rewriting the prefactor of Eq. (98) in terms of $\tau_{\text{Ly}\alpha}$, we can simplify Eq. (98) to

$$\Delta\dot{x}_{\downarrow}^{(\text{B})} = \frac{A_{2p,1s}}{\tau_{\text{Ly}\alpha}} x_{2p} [\Phi(-\infty) - 1], \quad (99)$$

where $\Phi(-\infty)$ is evaluated as a function of W . Note that in the $W \rightarrow 0$ limit, corresponding to negligible two-photon opacity $\Phi(-\infty) \rightarrow 1$, and this correction vanishes. We have compared Eq. (99) with the full numerical result for the $n \geq 3$ two-photon transitions with $\nu < \nu_{\text{Ly}\alpha}$ in Fig. 7.

C. Super- $\text{Ly}\alpha$ two-photon decays

We now turn to the super- $\text{Ly}\alpha$ two-photon decays from $n \geq 3$ levels of H I. These decays are fundamentally different from the sub- $\text{Ly}\alpha$ decays considered in the previous section. The reason is that a photon emitted at frequencies below $\text{Ly}\alpha$ has a nonzero probability of escaping the $\text{Ly}\alpha$ line, after which it is extremely unlikely to be reabsorbed; it may be Thomson or Rayleigh scattered, but it is unlikely to re-excite an H I atom and will consequently have little further impact on recombination. In contrast, all photons emitted above the $\text{Ly}\alpha$ frequency will eventually redshift down to $\text{Ly}\alpha$ and be absorbed. That is, for every super- $\text{Ly}\alpha$ two-photon decay there is an additional $\text{Ly}\alpha$ absorption, so the implied net number of decays to the H I $1s$ state is zero. For this reason, CS08 did not include these decays in their analysis. However, we can see from Fig. 3 that these decays do indeed result in a change in the recombination history. This is because of the *delay* between the two-photon decay that produces the photon at $\nu > \nu_{\text{Ly}\alpha}$ and the $\text{Ly}\alpha$ absorption. This section constructs an analytic approximation that accounts for the effect of this delay on recombination.

The key to our analytic approximation is to consider $x_+^{2\gamma}$, the number of nonthermal photons above the $\text{Ly}\alpha$ frequency per H atom produced by the two-photon decays. Initially, the photon spectrum is thermal, so $x_+^{2\gamma} = 0$. Long after recombination is over and all of the photons have redshifted below $\text{Ly}\alpha$, we also have $x_+^{2\gamma} = 0$. The net rate of decays to the ground state due to two-photon decays with $\nu > \nu_{\text{Ly}\alpha}$ and reabsorption of the emitted photons is

$$\dot{x}_{\downarrow}^{(\text{C})} = -\dot{x}_+^{2\gamma}. \quad (100)$$

The total number of such decays per H atom is

$$\int \dot{x}_{\downarrow}^{(\text{C})} dt = - \int \dot{x}_+^{2\gamma} dt = x_+^{2\gamma}(t_{\text{final}}) - x_+^{2\gamma}(t_{\text{init}}) = 0, \quad (101)$$

as expected.

It only remains to find $x_+^{2\gamma}(t)$. By mode counting, we find

$$x_+^{2\gamma} = \frac{8\pi}{c^3 n_{\text{H}}} \int_{\nu_{\text{Ly}\alpha}}^{\infty} \nu^2 \Delta f_{\nu} d\nu, \quad (102)$$

where Δf_ν is the nonthermal phase-space density of photons. Switching to the variable $y = h(\nu - \nu_{\text{Ly}\alpha})/T_r$, and taking the approximation that the distortion photons above $\text{Ly}\alpha$ are mostly in the regime $\nu/\nu_{\text{Ly}\alpha} - 1 \ll 1$,

$$x_+^{2\gamma} \approx \frac{8\pi\nu_{\text{Ly}\alpha}^2 T_r}{c^3 n_{\text{H}} h} \int_0^\infty \Delta f_\nu dy. \quad (103)$$

We may estimate the spectral distortion using Eq. (82). Rewriting it in terms of the distortion Δf_ν , we find

$$\frac{\Delta \dot{f}_\nu}{H\nu} = \frac{\partial \Delta f_\nu}{\partial \nu} - \kappa_\nu (\Delta f_\nu - \Delta f_\nu^0), \quad (104)$$

where

$$\Delta f_\nu^0 = f_\nu^0 - \frac{1}{e^{h\nu/T_r} - 1} \quad (105)$$

and the optical depth per unit frequency κ_ν is the same as that from Sec. VB. Under the same assumptions from Sec. VB that led to Eq. (85), we find

$$\Delta f_\nu^0 \approx \left(\frac{x_{2p}}{3x_{1s}} - e^{-h\nu_{\text{Ly}\alpha}/T_r} \right) e^{-h(\nu - \nu_{\text{Ly}\alpha})/T_r}. \quad (106)$$

We now assume the distortion is in steady state, i.e. that the left-hand side of Eq. (104) is zero. Defining

$$\Psi(\nu) = \left(\frac{x_{2p}}{3x_{1s}} - e^{-h\nu_{\text{Ly}\alpha}/T_r} \right)^{-1} \Delta f_\nu, \quad (107)$$

we convert Eq. (104) into

$$0 = \frac{d\Psi}{d\nu} - \kappa_\nu (\Psi - e^{-y}). \quad (108)$$

Using Eq. (87) for κ_ν , we convert this into

$$\frac{d\Psi}{dy} = W(e^y \Psi - 1). \quad (109)$$

This is the same equation as Eq. (93), except that now the interesting region is $0 < y < \infty$, and the “initial condition” is that $\Psi(\infty) = 0$. We define the integral

$$I = \int_0^\infty \Psi dy, \quad (110)$$

which is a function of W . Then Eq. (103) can be rewritten as

$$x_+^{2\gamma} \approx \frac{8\pi\nu_{\text{Ly}\alpha}^2 T_r}{c^3 n_{\text{H}} h} \left(\frac{x_{2p}}{3x_{1s}} - e^{-h\nu_{\text{Ly}\alpha}/T_r} \right) I. \quad (111)$$

The integral I as a function of W is plotted in Fig. 6. We evaluate it with an implicit ODE each time it is needed. From Eq. (109) we see that one should have $I \rightarrow 0$ as $W \rightarrow 0$, and $\Psi \rightarrow e^{-y}$, $I \rightarrow 1$ as $W \rightarrow \infty$. Figure 6 shows that this is precisely what happens.

Equation (100) presents an implementation difficulty: it depends on the derivative of $x_+^{2\gamma}$, which itself depends on \dot{x}_e because of the implicit dependence of x_{1s} and x_{2p} in

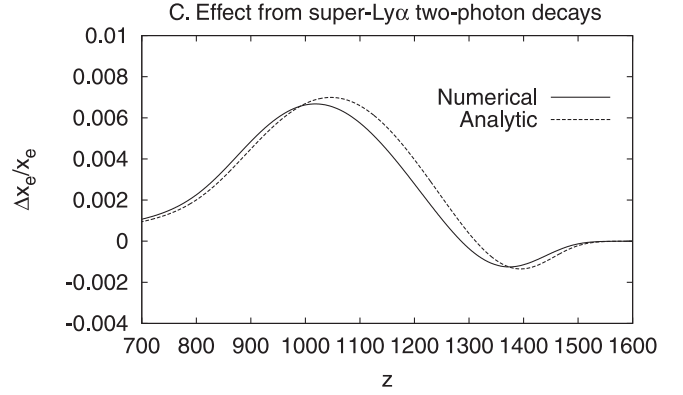


FIG. 8. A comparison of methods for handling the super- $\text{Ly}\alpha$ $n \geq 3$ two-photon decays. We show the numerical radiative transfer method (solid line) and the analytic approximation of Eqs. (100) and (111) (dashed line).

Eq. (111) on x_e . In principle, one could solve this by treating Eq. (100) as an implicit equation for $\dot{x}_+^{(C)}$. In practice, however, the effect of the two-photon decays from highly excited states is only a small perturbation on the recombination problem, so we have calculated Eq. (111) using the “standard” recombination history described in Sec. II. We then calculate $\dot{x}_+^{2\gamma}$ for use in Eq. (100) by numerically differentiating this precomputed table of $\dot{x}_+^{2\gamma}(t)$.

The result is shown in Fig. 8. One can see that the analytic approximation described here not only captures the qualitative fact that recombination is first accelerated and then delayed, because $x_+^{2\gamma}$ first increases and then decreases, but also captures the quantitative magnitude of the effect. The maximum of $x_+^{2\gamma}$ is 0.0061 at $z = 1360$, and one can see from Fig. 8 that this is indeed when the two-photon decays with $\nu > \nu_{\text{Ly}\alpha}$ switch from speeding up recombination to slowing it down.

The abundance $x_+^{2\gamma}$ of nonthermal photons above $\text{Ly}\alpha$ from two-photon decay using the analytic estimate is shown in Fig. 9.

D. Raman scattering

We now consider Raman scattering. Selection rules imply that Raman scattering to the $1s$ level is only possible from s and d levels. The most populated such level is $2s$, so we expect that Raman scattering is dominated by the $2s \leftrightarrow 1s$ process. This process contains contributions from two types of photons: the photons near the peak of the black-body spectrum, and the photons near $\text{H}\alpha$, which contribute resonant Raman scattering

$$\text{H}(2s) + h\nu' \rightarrow \text{H}(3p) \rightarrow \text{H}(1s) + h\nu, \quad (112)$$

which is resonant when $\nu' \approx \nu_{\text{H}\alpha}$ and $\nu \approx \nu_{\text{Ly}\beta}$. These pieces are (mostly) distinct: the radiation temperature is typically $0.02h\mathcal{R}$ ($z = 1200$), whereas the $\text{H}\alpha$ energy is

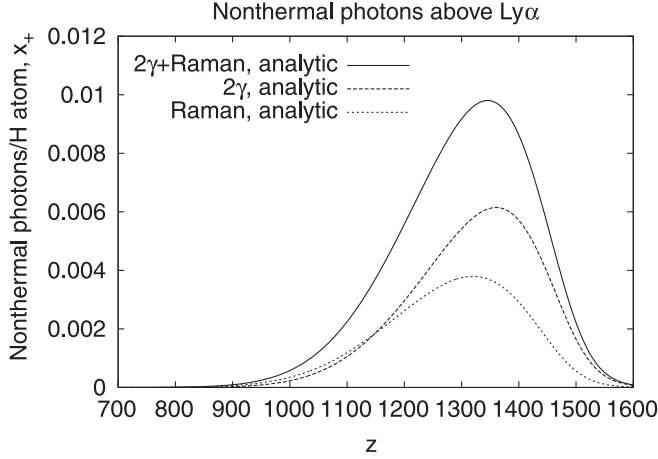


FIG. 9. The number of nonthermal photons above $\text{Ly}\alpha$ per H atom. The long-dashed curve shows the analytic estimate for $x_+^{2\gamma}$ (Sec. VC); the short-dashed curve shows the analytic estimate for x_+^R (Sec. VD); and the solid curve shows their sum $x_+^{2\gamma} + x_+^R$.

$0.14h\mathcal{R}$. Thus, the rate of Raman scattering as a function of frequency ν' of the incoming photon has two peaks—a nonresonant peak due to the large abundance of photons with energies of a few times T_r , and a resonant peak at $\text{H}\alpha$. There are additional peaks due to the other Balmer resonances $\text{H}\beta$, $\text{H}\gamma$, etc., but the abundance of these photons is at least a factor of a few less than $\text{H}\alpha$ (due to the Boltzmann factor), and the dipole matrix elements are smaller. This behavior can be seen in Fig. 1.

The two subclasses of Raman-scattering transitions—nonresonant and resonant ($\text{H}\alpha$)—are considered separately here.

I. Nonresonant scattering

In terms of effects on recombination, the nonresonant Raman scattering is similar to the super- $\text{Ly}\alpha$ two-photon decays in that a Raman scattering that produces an H I atom in the $1s$ state also produces a photon above the $\text{Ly}\alpha$ frequency. This photon will eventually redshift into the $\text{Ly}\alpha$ line and excite a hydrogen atom, so again, the net number of decays to the H I $1s$ state is zero. Again the importance for recombination derives from the delay between emission of the photon and its reabsorption. Thus, Eq. (100) applies to Raman scattering as well, except that in the case of Raman scattering we will need a formula for x_+^R , the number of nonthermal photons above $\text{Ly}\alpha$ contributed by nonresonant Raman scattering

$$\dot{x}_+^{(\text{D1})} = -\dot{x}_+^R. \quad (113)$$

Our challenge now is to find an equation for x_+^R . Reviewing the steps that lead to Eq. (108) shows that these are valid in the case of Raman scattering, except that we must replace the two-photon opacity with the Raman

opacity

$$\kappa_\nu = \frac{c^3 n_H}{8\pi H \nu^3} \sum_{nl} \left[\frac{g_{nl}}{g_{1s}} (1 + f_{\nu'}) x_{1s} - f_{\nu'} x_{2s} \right] \frac{dK_{nl}}{d\nu}, \quad (114)$$

in place of Eq. (83). Taking the $2s$ term to be dominant, recalling that $x_{2s} \ll x_{1s}$, and using the blackbody formula for $f_{\nu'}$, we can rewrite this as

$$\kappa_\nu \approx \frac{c^3 n_H}{8\pi H \nu^3} \frac{x_{1s}}{1 - e^{-y}} \frac{dK_{2s}}{d\nu}, \quad (115)$$

where $y = h\nu'/T_r = h(\nu - \nu_{\text{Ly}\alpha})/T_r$.

The Raman-scattering rate coefficient $dK_{2s}/d\nu$ can be rewritten using Eq. (B10) by noting that so long as $\nu' \ll \nu_{\text{H}\alpha}$, we may take $2p$ to be the dominant intermediate state. Using the matrix element $\langle 2s || \mathbf{r} || 2p \rangle^2 = 27a_0^2$

$$\frac{dK_{2s}}{d\nu} \approx \frac{\alpha_{\text{fs}}^6 \nu^3 \nu'}{\mathcal{R}^4 a_0^2} |\langle 1s || \mathbf{r} || 2p \rangle|^2. \quad (116)$$

The squared matrix element is related to the $\text{Ly}\alpha$ decay rate via the dipole emission formula

$$A_{\text{Ly}\alpha} = \frac{2}{9} \pi \alpha_{\text{fs}}^3 \frac{\nu_{\text{Ly}\alpha}^3}{\mathcal{R}^2 a_0^2} |\langle 1s || \mathbf{r} || 2p \rangle|^2, \quad (117)$$

from which we may derive

$$\frac{dK_{2s}}{d\nu} \approx \frac{9}{2\pi} A_{\text{Ly}\alpha} \left(\frac{\nu}{\nu_{\text{Ly}\alpha}} \right)^3 \alpha_{\text{fs}}^3 \frac{\nu'}{\mathcal{R}^2}. \quad (118)$$

This is just the lowest-order term in the power series expansion of $dK_{2s}/d\nu$; there are corrections of order ν'^2 , ν'^3 , etc. Because of the large prefactor of the next-lowest term we include it:

$$\frac{dK_{2s}}{d\nu} \approx \frac{9}{2\pi} A_{\text{Ly}\alpha} \left(\frac{\nu}{\nu_{\text{Ly}\alpha}} \right)^3 \alpha_{\text{fs}}^3 \frac{\nu'}{\mathcal{R}^2} \left(1 + 8.15 \frac{\nu'}{\mathcal{R}} \right). \quad (119)$$

The coefficient 8.15 was determined by numerical differentiation of $dK_{2s}/d\nu$. It is positive because of constructive interference between the dominant $2s \rightarrow 2p \rightarrow 1s$ Raman-scattering pathway and the neighboring $2s \rightarrow 3p \rightarrow 1s$ pathway.

Making this approximation and plugging into Eq. (115), we find

$$\kappa_\nu \approx \frac{9c^3 n_H A_{\text{Ly}\alpha}}{16\pi^2 H \nu_{\text{Ly}\alpha}^3} \frac{x_{1s}}{1 - e^{-y}} \alpha_{\text{fs}}^3 \left(1 + 8.15 \frac{\nu'}{\mathcal{R}} \right) \frac{\nu'}{\mathcal{R}^2}. \quad (120)$$

We may rewrite this using the $\text{Ly}\alpha$ optical depth, Eq. (90), as

$$\kappa_\nu \approx \frac{3}{2\pi} \alpha_{\text{fs}}^3 \tau_{\text{Ly}\alpha} \left(\frac{T_r}{h\mathcal{R}} \right)^2 \left(1 + 8.15 \frac{T_r}{h\mathcal{R}} y \right) \frac{\nu'}{1 - e^{-y}}. \quad (121)$$

The Raman-scattering version of Eq. (108) is then

$$\frac{d\Psi}{dy} = V \left(1 + 8.15 \frac{T_r}{h\mathcal{R}} y \right) \frac{y}{1 - e^{-y}} (\Psi - e^{-y}), \quad (122)$$

where

$$V = \frac{3}{2\pi} \alpha_{\text{fs}}^3 \tau_{\text{Ly}\alpha} \left(\frac{T_r}{h\mathcal{R}} \right)^2. \quad (123)$$

Equation (122) has the integral solution,

$$\Psi(y) = V \int_y^\infty \frac{y' e^{-y'}}{1 - e^{-y'}} \left(1 + 8.15 \frac{T_r}{h\mathcal{R}} y' \right) e^{-\tau(y, y')} dy', \quad (124)$$

where

$$\tau(y, y') = V \int_y^{y'} \frac{y''}{1 - e^{-y''}} \left(1 + 8.15 \frac{T_r}{h\mathcal{R}} y'' \right) dy'' \quad (125)$$

is the optical depth between y and y' . We may then write the dimensionless integral of the spectral distortion

$$\begin{aligned} \mathcal{J} &\equiv \int_0^\infty \Psi(y) dy \\ &= V \int_0^\infty dy' \frac{y' e^{-y'}}{1 - e^{-y'}} \left(1 + 8.15 \frac{T_r}{h\mathcal{R}} y' \right) \\ &\quad \times \int_0^{y'} dy e^{-\tau(y, y')}. \end{aligned} \quad (126)$$

If τ were negligible, then the inner integral would collapse to give y' , and the outer integral would be dominated by the region $y' \sim 2$. In practice, for $V \ll 1$ (the maximum is 0.046 at $z = 1184$) τ will be $\ll 1$ in this region; for much larger y' it may be significant, but all it does is suppress a portion of the integral that is already negligible. Thus, we may neglect τ and replace the integral $\int_0^{y'} dy e^{-\tau(y, y')}$ with y' . This gives

$$\begin{aligned} \mathcal{J} &\approx V \int_0^\infty dy' \frac{y'^2 e^{-y'}}{1 - e^{-y'}} \left(1 + 8.15 \frac{T_r}{h\mathcal{R}} y' \right) \\ &= \left[2\zeta(3) + 8.15 \frac{T_r}{h\mathcal{R}} \frac{\pi^4}{15} \right] V \\ &= 2\zeta(3) V \left(1 + 22.0 \frac{T_r}{h\mathcal{R}} \right). \end{aligned} \quad (127)$$

We may now assemble all of the pieces and plug them into Eq. (102): the spectral distortion is given in terms of Ψ by Eq. (107), and the integral of Ψ is Eq. (127). The normalization V is given by Eq. (123). Combining everything gives

$$\begin{aligned} x_+^R &= \frac{27}{2} \zeta(3) \left(\frac{T_r}{hc} \right)^3 \frac{\tau_{\text{Ly}\alpha}}{n_{\text{H}}} \left(1 + 22.0 \frac{T_r}{h\mathcal{R}} \right) \\ &\quad \times \left(\frac{x_{2p}}{3x_{1s}} - e^{-h\nu_{\text{Ly}\alpha}/T_r} \right). \end{aligned} \quad (128)$$

The same implementation difficulty arises with Eq. (113) as arose with Eq. (100), namely, that in order

to compute $\Delta\dot{x}_\downarrow^{(\text{D})}$ one needs to know x_+^R which in turn requires one to already know the recombination history. We solve this problem by taking x_+^R from the standard recombination history (Sec. II), just as we did in Sec. V C.

2. H α resonance

Resonant Raman scattering via the $3p$ intermediate state [Eq. (112)] will induce a red damping wing correction to the Ly β escape probability similar to the corrections discussed in Sec. V B for the two-photon corrections to Ly α . Precisely the same calculation applies here except that the line width, Eq. (91) gets replaced by

$$\bar{W}_\beta \approx \frac{\tau_{\text{Ly}\beta}}{4\pi^2} \frac{A_{3p,2s}}{1 - e^{-h\nu_{3p,2s}/T_r}}. \quad (129)$$

[The only differences in the calculation are the replacement of the initial state $2s$ instead of nl ; the intermediate state $3p$ instead of $2p$; the use of the Raman-scattering rate coefficient $dK/d\nu$ in place of $d\Lambda/d\nu$; since the low-frequency photon is in the initial state the use of $1 + f_{\nu'}$ in Eq. (114) instead of $f_{\nu'}$, which appears in Eq. (83); and no factor of $(2l + 1)/3$ because the Einstein coefficient $A_{3p,2s}$ in Eq. (129) has the states reversed from $A_{nl,2p}$ in Eq. (91)]. Once again we can construct a dimensionless variable $W_\beta = h\bar{W}_\beta/T_r$ and introduce a correction

$$\Delta\dot{x}_\downarrow^{(\text{D2})} = \frac{A_{3p,1s}}{\tau_{\text{Ly}\beta}} x_{3p} [\Phi(-\infty) - 1], \quad (130)$$

where $\Phi(-\infty)$ is evaluated from Eq. (93) using W_β instead of W .

The parameter W_β measures the normalization of the Raman-scattering opacity [to the Ly β photon, not the H α photon; i.e. for the reverse reaction of Eq. (112)], just as W measured the normalization of the two-photon opacity to photons near Ly α . We find that W_β reaches a maximum of 0.94 at $z = 981$. It is much larger than the corresponding W for the Ly α line, because two-photon absorption of a photon near Ly α requires a three-particle interaction between the Ly α photon, an atom, and a rare Balmer photon in the Wien tail of the CMB; whereas a photon near Ly β can be Raman-scattered in a two-body interaction without the help of any initial-state Balmer photons. The larger W_β means that the corrections to Ly β are of similar overall importance to the corrections to Ly α , even though in an absolute sense Ly β is less important.

The overall change in the downward decay rate from Raman scattering is the sum of the nonresonant and resonant contributions

$$\Delta\dot{x}_\downarrow^{(\text{D})} = \Delta\dot{x}_\downarrow^{(\text{D1})} + \Delta\dot{x}_\downarrow^{(\text{D2})}. \quad (131)$$

However, it should be noted that the downward rate D2 [Eq. (130)] is emitted in the red damping wing of Ly β and is added to the Ly β spectral distortion $f_{\text{Ly}\beta-} \equiv f_{3p,1s-}$ appearing on the left-hand side of Eq. (6). It is also in-

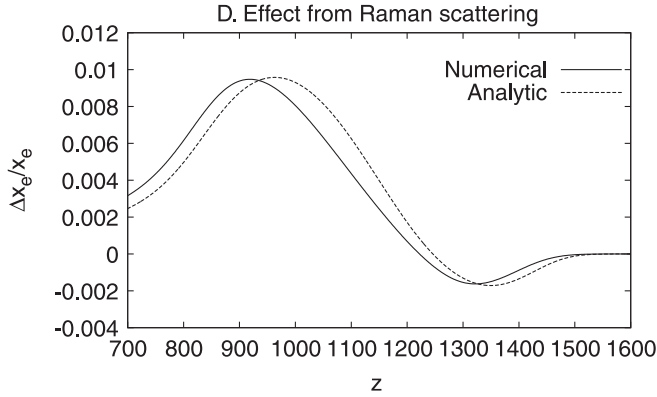


FIG. 10. A comparison of methods for handling Raman scattering. We show the numerical radiative transfer method (solid line) and the analytic approximation of Eq. (131) and (128) (dashed line).

cluded as an increase in the $3p \rightarrow 1s$ transition rate $\dot{x}_{3p \rightarrow 1s}$. It therefore feeds back at later times on $\text{Ly}\alpha$. In contrast, the rate D1 has all radiative transfer effects already included, and it is included along with Eq. (100) in the $2p \rightarrow 1s$ rate. (It is not obvious whether it is better to analytically approximate D1 as being a net rate from the $2s$ or $2p$ state, since the Raman scattering occurs from $2s$ but re-excitations are to $2p$. In practice, $2s$ and $2p$ are in equilibrium in the redshift range considered so this choice does not matter.)

The effect of this correction on recombination is shown in Fig. 10. Our analytic approximation reproduces most of the features of the full numerical result. The main deficiency is that the effects on $\Delta x_e/x_e$ occur earlier in the analytic calculation than in the numerical result by about $\Delta z \sim 50$. This error might be due to the approximation that x_+^R instantaneously reaches its steady-state value, but we defer investigation of this possibility to future work.

As a summary, we show in Fig. 11 the total error in using the analytical approximations presented here in place of more accurate numerical two-photon radiative transfer. The analytic approximations have reduced the maximal fractional error in x_e from 1.3% with no treatment of the

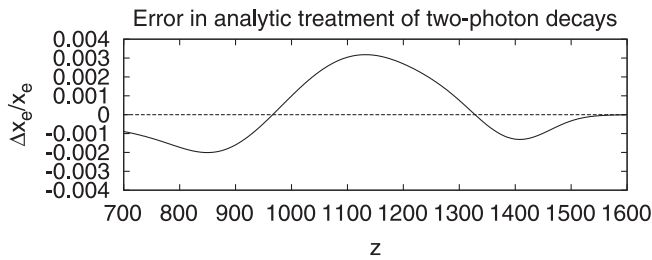


FIG. 11. The error in using analytic approximations for the two-photon effects. The vertical axis plots $\Delta x_e/x_e$, with the sign convention positive where the analytic approximation overestimates x_e and negative where it underestimates x_e .

two-photon effects to 0.3%, the maximal difference between the analytic and numerical treatments.

VI. EFFECT ON THE CMB

We have calculated the effect of our corrected recombination history on the CMB anisotropies using the CMBFAST [45] Boltzmann code. The results are shown in Fig. 12. For the temperature and E -mode polarization, we show the fractional change in the power spectrum C_ℓ^{TT} or C_ℓ^{EE} . For the cross spectrum we show the change in correlation coefficient, defined by $\rho_\ell^{TE} = C_\ell^{TE}/(C_\ell^{TT}C_\ell^{EE})^{1/2}$. This is more useful than $\Delta C_\ell^{TE}/C_\ell^{TE}$, which is ill-behaved when C_ℓ^{TE} passes through zero. For comparison, we show the analytic approximations of Sec. V. (Note that we denote the CMB multipoles by ℓ to avoid confusion with atomic orbital angular momentum l .)

While the mapping from the recombination history to the CMB power spectrum is in general complicated [9,11], the qualitative features of Fig. 12 are easily understood. The major effect on the temperature is that the decrease in

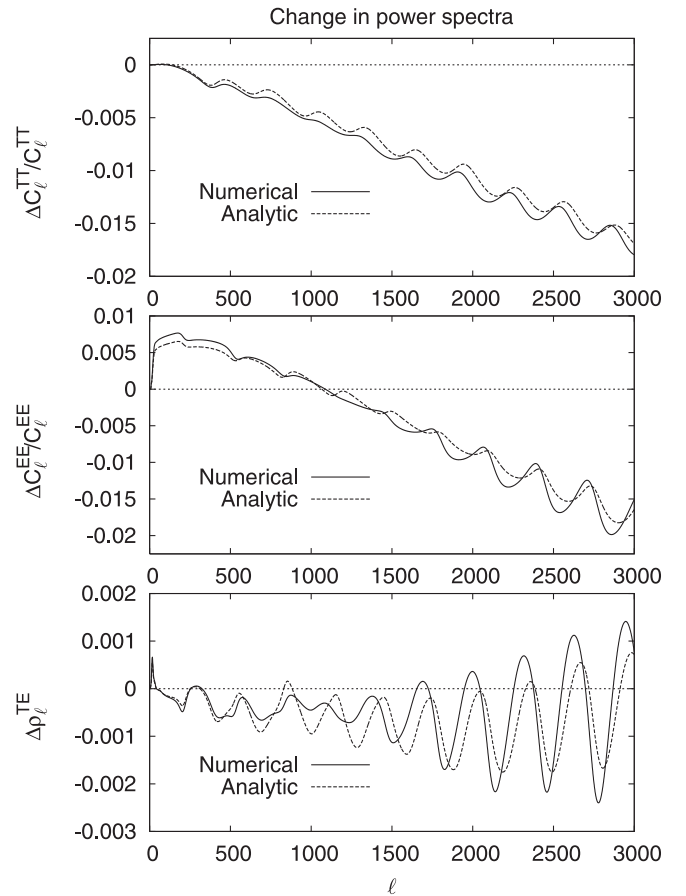


FIG. 12. The change in CMB temperature power spectrum, E -mode polarization, and cross correlation due to the two-photon processes described in this paper. The full calculation is shown with the solid lines, and the prediction from our simple analytic model is shown with the dashed lines.

free electron abundance in the $1100 < z < 1500$ range increases the mean free path of the photons and hence the Silk damping length [46–49]. This means that the oscillations of the baryon-photon fluid are damped more effectively when two-photon transitions are included than in the standard picture, and hence the power spectrum is decreased. This effect is most significant at the higher multipoles where damping is most effective. In principle, changes in the recombination history could also have moved the surface of last scatter, thereby shifting the acoustic peak positions and leading to oscillations in $\Delta C_\ell^{TT}/C_\ell^{TT}$; however, as one can see from the figure, this is a small effect. This is because the accelerating effect on recombination of sub-Ly α two-photon decays and the decelerating effect of super-Ly α decays and Raman scattering cancel at $z \approx 1100$. Thus, purely by accident, the surface of last scatter is not significantly shifted.

Silk damping also suppresses the E -mode polarization, and again this suppression is evident in the figure at high multipoles. However, this is not the only effect of the new recombination history. In our calculation, recombination is proceeding more slowly at the surface of last scatter, i.e. x_e is decreasing more slowly than in the standard picture, so the surface of last scatter is broader. Since generation of polarization requires a sufficiently low Thomson opacity to generate a local photon quadrupole by free streaming, and sufficiently high Thomson opacity to convert this quadrupole into polarization, the broader surface of last scatter enhances the E -mode polarization [49]. Since Silk damping increases rapidly as one goes to small scales, the broadening of the surface of last scatter wins at low ℓ ($\ell \lesssim 1000$), but Silk damping wins at high ℓ , leading to the characteristic “positive, then negative” shape seen in the middle panel of Fig. 12. At very low multipoles $\ell \lesssim 50$ the old and new calculations agree despite the broader surface of last scatter; this is because the large-scale polarization is dominated by scattering at reionization, which we have left unchanged. However, the small changes described here are unobservable at such low ℓ because of cosmic variance.

The temperature-polarization cross-correlation coefficient is only slightly affected ($\lesssim 0.002$) because to a first approximation the Silk damping and the broadening of the last scattering surface only give a wavenumber-dependent rescaling of the T and E . The only recombination process that would substantially change ρ_ℓ^{TE} would be to move the surface of last scatter, and as noted above, our corrections to recombination do not significantly move the surface of last scatter.

One can estimate the importance of the changes in the recombination history for various CMB experiments by computing

$$Z^2 = \sum_{\ell\ell'} F_{\ell\ell'} \Delta C_\ell^{TT} \Delta C_{\ell'}^{TT}, \quad (132)$$

where $F_{\ell\ell'}$ is the Fisher matrix for the temperature power

spectra. If all cosmological parameters were known perfectly and there were no systematic errors then Z represents the number of sigma at which the changes in the recombination history could be detected. In the realistic case in which one is trying to measure the cosmological parameters, then the maximum possible bias from using the incorrect recombination history is Z sigma; the actual bias will depend on whether the parameter in question produces a similar or orthogonal change in the C_ℓ 's to the change from the recombination history. For the WMAP 5-year Fisher matrix $F_{\ell\ell'}$ [2,50] we find $Z = 0.40$, i.e. a 0.40σ effect for statistical errors only. In practice, the calibration, beam, and point source uncertainties are significant in the range $200 \lesssim \ell \lesssim 800$ relevant for WMAP [51–54]; if one includes them by adding these modes to the inverse Fisher matrix, this is reduced to 0.27σ . A similar calculation is possible for the Arcminute Cosmology Bolometer Array Receiver final data release [55]; here one finds a 0.48σ effect for the Fisher matrix only, which is reduced to 0.19σ after marginalizing over the beam and calibration uncertainty. Thus, the corrections described here have negligible impact on the interpretation of the currently published WMAP and Arcminute Cosmology Bolometer Array Receiver data [2,55]. However, for the upcoming Planck mission [56], if we use the data model from the Dark Energy Task Force [57] including temperature anisotropies in the 100, 143, and 217 GHz bands, and E -mode polarization in the 143 and 217 GHz bands [including C_ℓ^{EE} in the sum and Fisher matrix in Eq. (132)], we find $Z = 7$, i.e. a 7σ effect.

We have also calculated the importance of the error using the analytic approximation, i.e. we obtain Z^2 from Eq. (132) but using the difference between numerical and analytic results for ΔC_ℓ^{TT} . This gives $Z = 0.8$ for Planck, implying that the analytic treatment of two-photon decays in this paper is sufficient to reduce the errors from 7σ to 0.8σ for Planck. While this is an enormous improvement, one would like to do better in the future, perhaps by computing the finite lag time required for $x_+^{2\gamma}$ to reach its steady-state value, or (less attractively) with a fudge factor fit to the full numerical result.

In addition to the CMB anisotropies, one could also consider the global spectral distortion from recombination. While observation of this signal will be challenging due to foregrounds [16] and the sensitivity requirements [58], it has prompted a large amount of work over the years because it is a direct probe of recombination physics [12,13,16,59–62]. Our analysis has shown a ~ 1 order of magnitude increase in the spectral distortion blueward of Ly α ; however, all of these photons are absorbed and reprocessed when they redshift into the Ly α line, and in our calculation the effect on the observable spectral distortion (red side of Ly α) is at the level of a few percent. We have not followed the lower-frequency distortion due to Balmer, Paschen, etc. lines in this paper, but the excited

level occupations typically change by a few percent or less and we expect the final spectral distortion due to these lines to change by a similar amount. Given that the spectral distortion has not yet been detected, it seems unlikely that the corrections considered in this paper will be important for spectral distortion studies in the foreseeable future.

VII. CONCLUSION

In this paper, we have reconsidered the physics of two-photon decays in hydrogen recombination, with particular emphasis on decays from highly excited levels $n \geq 3$ of H I. We reviewed the conceptual difficulties that have plagued earlier attempts to incorporate these transitions, particularly those that arise from sequential “ $1 + 1$ ” decays such as $3d \rightarrow 2p \rightarrow 1s$. We argued that the notion of an effective two-photon decay coefficient, useful and beloved for the $2s \rightarrow 1s$ decay of H I, must be abandoned for the higher states in order to resolve these conceptual difficulties. Rather, one must resolve these two-photon decays as a function of frequency of the emitted photons and solve the full radiative transfer problem for two-photon emission and absorption.

In addition to the two-photon decays from the highly excited states, we have also considered Raman scattering and two-photon recombination. These are physically and mathematically very similar processes, with the differences being moving one photon from the final to initial state (for Raman scattering) and moving the initial electron from a bound to a free state (for two-photon recombination). Neither process has previously been investigated in the cosmological context.

Our major phenomenological findings are as follows:

- (a) The traditional treatment of two-photon decays from $2s$ does not account for stimulated emission and reabsorption of the $\text{Ly}\alpha$ spectral distortion. The maximal correction to x_e from these effects is 0.6% and over most of the redshift range the net effect is to delay recombination. We agree with both the physical picture and the calculation of past studies of these effects [23,24].
- (b) Recombination is sped up by sub- $\text{Ly}\alpha$ two-photon decays from higher excited states; the most important initial state is $3d$. Physically, inclusion of this effect involves a combination of two-photon emission and absorption, which must be handled through a radiative transfer calculation; there is no relevant effective two-photon decay coefficient Λ_{3d} analogous to that for $2s$. We find a maximum change of 1.7% in x_e .
- (c) Super- $\text{Ly}\alpha$ two-photon decays initially speed up recombination by supplying another route to the ground state of hydrogen. These photons produce a delayed feedback that slows down recombination at $z < 1100$ as they redshift down to $\text{Ly}\alpha$. The maxi-

um effect on x_e is 0.7%. The time dependence of the line profile plays a central role in this effect.

- (d) CMB photons can Raman scatter from hydrogen atoms in the $2s$ level and insert the atom into the ground state. Again this initially speeds up recombination but produces photons blueward of $\text{Ly}\alpha$ that produce delayed feedback. The maximum effect on x_e is 1.0%.
- (e) Direct two-photon recombinations to the ground state of hydrogen are negligible.
- (f) The net effect of the two-photon transitions on the CMB power spectrum is twofold: (i) they suppress the temperature and polarization power spectra by up to 2% at $\ell \sim 3000$ due to increased Silk damping from the faster early stages of recombination; and (ii) they increase the degree-scale polarization by $\sim 0.7\%$ due to the wider surface of last scatter. At “Fisher matrix” level the net effect is 7σ for Planck.

We have constructed analytic approximations to all of the important two-photon processes that reproduce qualitatively and quantitatively the numerical radiative transfer results. The error from using these analytic approximations instead of doing full two-photon radiative transfer will be 0.8σ for Planck. This is reassuring: we anticipate that once Planck identifies the relevant portion of parameter space, this $\sim 0.8\sigma$ correction to $x_e(z)$ can be added as a fudge factor. However, much work lies ahead in order to make a full analysis of Planck data computationally feasible: our code even with the analytic approximations still takes ~ 1 day to run on a desktop machine, due primarily to the small step size of $\Delta \ln a = 4.25 \times 10^{-5}$. We chose this time step because of our method of following feedback: the excitation in the $1s - np$ line requires us to look up the photon phase-space density at a slightly earlier time on the red side of the $1s - (n+1)p$ line, and the time step must be small enough that the interpolation table for $f_{1s,(n+1)p-}$ has already been built. Iterative feedback methods, such as those used in Ref. [26], can use much larger time steps and hence should be better suited to fast calculation.

In the future, we plan to incorporate these developments into a more complete hydrogen recombination model that would include effects not considered here. Some of these effects, such as increasing the number of shells n_{max} and collisions, have previously received treatment in the literature [25], but ultimately one must *simultaneously* incorporate all important processes. Another process that may interact with some of the effects here is $\text{Ly}\alpha$ diffusion and recoil, for which one would have to add Fokker-Planck terms to the radiative transfer equations described herein. There is some work on this problem [29,36–38] and for the analogous problem in the He I 594 Å line [26]; however, given the importance of two-photon transitions, their optically thick nature, and the $>1\%$ correction from effects that owe their existence to the time dependence of the line profile, it is clear that we need to solve the full

time-dependent radiative problem including both Ly α diffusion and two-photon opacity. This will be the subject of future work. A final goal will be to speed up the recombination calculation to produce a recombination code suitable for inclusion into Markov chain Monte Carlo codes for cosmological parameter estimation.

ACKNOWLEDGMENTS

C. H. thanks Eric Switzer for a careful reading of the draft of this paper; and Jens Chluba, Rashid Sunyaev, Dan Grin, and Ulrich Jentschura for illuminating discussions about the physics. This paper was improved by the suggestions from the referee, Douglas Scott. C. H. is supported by DOE Contract No. DE-FG03-92-ER40701.

APPENDIX A: MATTER TEMPERATURE

This appendix constructs a solution for the matter temperature evolution valid at early times. We can write a formal integral solution to Eq. (15), since the equation is (if we fix x_e) a first-order linear ODE in T_m/T_r . The usual procedure gives

$$\frac{T_m}{T_r}(t) = - \int_0^t \frac{J(\tilde{t})}{1 + J(\tilde{t})} e^{-K(\tilde{t})} K'(\tilde{t}) d\tilde{t}, \quad (\text{A1})$$

where

$$J = \frac{8x_e \sigma_T a_r T_r^4}{3(1 + f_{\text{He}} + x_e) m_e c H} \quad (\text{A2})$$

and K is the integral defined by

$$K(\tilde{t}) = \int_{\tilde{t}}^t H(\tilde{t}) [1 + J(\tilde{t})] d\tilde{t}, \quad (\text{A3})$$

with $K(t) = 0$ and $K'(\tilde{t}) < 0$. The initial value or “particular solution” is unimportant since $K(0) = \infty$. Since $\int_0^t e^{-K(\tilde{t})} K'(\tilde{t}) d\tilde{t} = -1$, we may rewrite Eq. (A1) as

$$\frac{T_m}{T_r}(t) = 1 + \int_0^t \frac{1}{1 + J(\tilde{t})} e^{-K(\tilde{t})} K'(\tilde{t}) d\tilde{t}. \quad (\text{A4})$$

This is still only a formal solution, because we cannot in general solve for x_e without knowing T_m . However, in the case $J \gg 1$, which occurs during recombination (but is violated at $z \lesssim 500$), then the exponential factor $e^{-K(\tilde{t})}$ kills the integrand in Eq. (A4) unless $\tilde{t} \approx t$. In this case, we may substitute $J(\tilde{t}) \approx J(t)$, and for $J \gg 1$ we get

$$\frac{T_m}{T_r}(t) = 1 - \frac{1}{J(t)}, \quad (\text{A5})$$

i.e. Eq. (25).

APPENDIX B: TWO-PHOTON TRANSITION RATES

This appendix is concerned with calculating the transition rates for spontaneous two-photon decay ($nl \rightarrow 1s$),

Raman scattering ($nl \rightarrow 1s$), and spontaneous two-photon recombination to the ground state of hydrogen. Stimulated rates and reverse reactions ($1s \rightarrow nl$ and two-photon ionization) are obtainable by insertion of the usual $1 + f_\nu$ factor and by the principle of detailed balance, respectively. Note that the two-photon selection rules allow these reactions only when the initial state has $l = 0$ or $l = 2$ (see DG05). This also holds for two-photon recombination, although in this case, it is acceptable to sum the two cases, because for the $H^+ + e^-$ continuum the l sublevels are statistically populated.

We will have to consider both discrete and continuous states here. The latter appear in summations over intermediate states in two-photon decay and Raman scattering, and as initial states in two-photon recombination. They are characterized by an imaginary principal quantum number to agree with the usual formula $E_e = -\mathcal{R}/n^2$ (we take the $\Im n > 0$ branch), or alternatively by their wave number

$$k = \frac{\sqrt{2\mu E_e}}{\hbar} = i \frac{\sqrt{\mu}}{\hbar n}. \quad (\text{B1})$$

The conventional normalization of the unbound states is that the radial wavefunction $R(r)$ [with $r\psi(r, \theta, \phi) = R(r)Y_{lm}(\theta, \phi)$] is chosen to oscillate between -1 and $+1$, i.e. we normalize to a sphere of radius $X = 2$. In this case, continuum states are formally separated in wave number by $\Delta k = \pi/X = \pi/2$. Therefore, in all calculations of summations over intermediate p states, one should in actual computation make the replacement

$$\sum_{\text{all } n} \rightarrow \sum_{n=2}^{\infty} + \frac{2}{\pi} \int_0^{\infty} dk, \quad (\text{B2})$$

where we will write “all n ” for emphasis and $n = ik/(\mu^{1/2}\hbar)$ for continuum states.

1. Two-photon decay

The spontaneous two-photon decay rate for the process

$$H(nl) \rightarrow H(1s) + h\nu + h\nu' \quad (\text{B3})$$

is given by

$$\frac{d\Lambda}{d\nu} = \frac{\alpha_{\text{fs}}^6 \nu^3 \nu'^3}{108(2l+1)\mathcal{R}^6} |\mathcal{M}|^2, \quad (\text{B4})$$

where $\alpha_{\text{fs}} \approx 1/137$ is the fine structure constant, \mathcal{R} is the hydrogen Rydberg constant, and the matrix element is

$$\mathcal{M} = \frac{2}{a_0^2} \sum_{\text{all } N} \langle nl || \mathbf{r} || Np \rangle \langle Np || \mathbf{r} || 1s \rangle \left(\frac{1}{1 - N^{-2} - \nu/\mathcal{R}} + \frac{1}{1 - N^{-2} - \nu'/\mathcal{R}} \right). \quad (\text{B5})$$

(See Eq. 13 of Ref. [22]). The summation over intermediate states is carried out over both the discrete values of N and the continuous spectrum ($N^{-2} < 0$, restricted of course

to p symmetry). Note that \mathcal{M} as defined here is dimensionless, and hence so is $d\Lambda/d\nu$.

One can write the total two-photon rate

$$\Lambda_{nl} = \int_{(1-n^{-2})\mathcal{R}/2}^{(1-n^{-2})\mathcal{R}} \frac{d\Lambda}{d\nu} d\nu. \quad (\text{B6})$$

The cutoff at half the maximum frequency avoids double counting, since one photon is always above this cut and the other below.

2. Raman scattering

Now, we are interested in the reaction

$$\text{H}(nl) + h\nu' \rightarrow \text{H}(1s) + h\nu. \quad (\text{B7})$$

This reaction is related to the two-photon decay rate by crossing symmetry. Therefore, whereas the two-photon decay rate was

$$\frac{d\Gamma_{nl,1s}}{d\nu}(2\gamma) = \frac{\alpha_{\text{fs}}^6 \nu^3 \nu'^3}{108(2l+1)\mathcal{R}^6} |\mathcal{M}|^2 (1+f_\nu)(1+f_{\nu'}), \quad (\text{B8})$$

the Raman-scattering rate is

$$\frac{d\Gamma_{nl,1s}}{d\nu}(\text{Raman}) = \frac{\alpha_{\text{fs}}^6 \nu^3 \nu'^3}{108(2l+1)\mathcal{R}^6} |\mathcal{M}|^2 (1+f_\nu)f_{\nu'}, \quad (\text{B9})$$

where the matrix element \mathcal{M} is given by Eq. (B5). It is convenient for us to define the coefficient

$$\frac{dK_{nl}}{d\nu} = \frac{\alpha_{\text{fs}}^6 \nu^3 \nu'^3}{108(2l+1)\mathcal{R}^6} |\mathcal{M}|^2 \quad (\text{B10})$$

that contains the matrix element and kinematic factors but not the state of the radiation field. Then

$$\frac{d\Gamma_{nl,1s}}{d\nu}(\text{Raman}) = \frac{dK_{nl}}{d\nu} (1+f_\nu)f_{\nu'}. \quad (\text{B11})$$

3. Two-photon recombination

We also require the rate coefficient for two-photon recombination

$$\text{H}^+ + e^- \rightarrow \text{H}(1s) + h\nu + h\nu', \quad (\text{B12})$$

where the initial-state electron has kinetic energy E_e , and the final-state photons satisfy $h\nu + h\nu' = E_e - E_{1s}$. We define the coefficient $\alpha^{(2)}(\nu, \nu') n_e n_p$ as the number of recombinations per unit time per unit frequency ($d\nu$) in a

gas of monoenergetic electrons and stationary protons, and in the absence of a stimulating radiation field. The computation of this rate is similar to that of the two-photon decay, except that now the initial state is in the continuum ($n^2 < 0$). This coefficient is a modification of Eq. (B4):

$$\alpha^{(2)}(\nu, \nu') = \sum_{l=0,2} \bar{N}_e(E, l) \frac{\alpha_{\text{fs}}^6 \nu^3 \nu'^3}{108(2l+1)\mathcal{R}^6} |\mathcal{M}|^2, \quad (\text{B13})$$

where $\bar{N}_e(E, l)$ is the mean number of electrons per normalization volume with angular momentum l , assuming unit electron density. This is given by

$$\bar{N}_e(E, l) = 2\pi(2l+1) \frac{X}{k^2} = \pi(2l+1) \frac{\hbar^2 X}{\mu E_e}, \quad (\text{B14})$$

where X is the radius of the normalization sphere. (This is most easily proved by considering the decomposition of a plane wave with unit probability density, $e^{i\mathbf{k}\cdot\mathbf{r}}$ into spherical harmonics, integrating the l th partial wave out to radius X and taking the large- X limit $\int_0^X r^2 |j_l(kr)|^2 dr \rightarrow X/2k^2$.) This normalization volume cancels the $X^{-1/2}$ dependence of the matrix element \mathcal{M} from the initial-state wave function. For our choice of normalization, $X = 2$ and we have

$$\alpha^{(2)}(\nu, \nu') = \sum_{l=0,2} \frac{\pi \alpha_{\text{fs}}^6 \hbar^2 \nu^3 \nu'^3}{54 \mathcal{R}^6 \mu E_e} |\mathcal{M}|^2. \quad (\text{B15})$$

The matrix element \mathcal{M} is given by Eq. (B5). Once again we restrict $\nu > \nu'$. Note that the normalization of the free state forces $\psi(\mathbf{r})$ to have units of cm^{-1} instead of $\text{cm}^{-3/2}$, so for two-photon recombination \mathcal{M} has units of $\text{cm}^{1/2}$. Thus, $\alpha^{(2)}(\nu, \nu')$ has units of cm^3 .

4. Computation of matrix elements

The two-photon matrix elements \mathcal{M} can in principle be computed by brute force from Eq. (B5). This was the approach taken in CS08 for hydrogen and in Hirata and Switzer [22] for helium. Because of the cancellations among the various intermediate states for large n , however, we have chosen instead to use the Green function method, which is expected to be numerically stable even for very large n . The Green function method is described in the recombination context by Eqs. (24–26) of Ref. [22], and more generally by Refs. [63,64].

-
- [1] D.N. Spergel *et al.*, *Astrophys. J. Suppl. Ser.* **170**, 377 (2007).
 - [2] E. Komatsu *et al.*, arXiv:0803.0547.
 - [3] W. Hu, *Phys. Rev. D* **65**, 023003 (2001).
 - [4] U. Seljak, N. Sugiyama, M. White, and M. Zaldarriaga, *Phys. Rev. D* **68**, 083507 (2003).

- [5] P.J.E. Peebles, *Astrophys. J.* **153**, 1 (1968).
- [6] Y.B. Zeldovich, V.G. Kurt, and R.A. Sunyaev, *Zh. Eksp. Teor. Fiz.* **55**, 278 (1968).
- [7] S. Seager, D.D. Sasselov, and D. Scott, *Astrophys. J. Lett.* **523**, L1 (1999).
- [8] S. Seager, D.D. Sasselov, and D. Scott, *Astrophys. J.*

- Suppl. Ser. **128**, 407 (2000).
- [9] W. Y. Wong, A. Moss, and D. Scott, *Mon. Not. R. Astron. Soc.* **386**, 1023 (2008).
- [10] V. V. Sobolev, *Moving Envelopes of Stars* (Harvard University Press, Cambridge, England, 1960).
- [11] A. Lewis, J. Weller, and R. Battye, *Mon. Not. R. Astron. Soc.* **373**, 561 (2006).
- [12] J. A. Rubino-Martín, J. Chluba, and R. A. Sunyaev, *Mon. Not. R. Astron. Soc.* **371**, 1939 (2006).
- [13] J. Chluba and R. A. Sunyaev, *Astron. Astrophys.* **458**, L29 (2006).
- [14] V. K. Dubrovich and N. N. Shakhvorostova, *Astron. Lett.* **30**, 509 (2004).
- [15] V. K. Dubrovich and S. I. Grachev, *Astron. Lett.* **30**, 657 (2004).
- [16] W. Y. Wong, S. Seager, and D. Scott, *Mon. Not. R. Astron. Soc.* **367**, 1666 (2006).
- [17] J. Chluba and R. A. Sunyaev, *Astron. Astrophys.* **478**, L27 (2008).
- [18] R. A. Sunyaev and J. Chluba, arXiv:0802.0772.
- [19] V. K. Dubrovich and S. I. Grachev, *Astron. Lett.* **31**, 359 (2005).
- [20] W. Y. Wong and D. Scott, *Mon. Not. R. Astron. Soc.* **375**, 1441 (2007).
- [21] J. Chluba and R. A. Sunyaev, *Astron. Astrophys.* **480**, 629 (2008).
- [22] C. M. Hirata and E. R. Switzer, *Phys. Rev. D* **77**, 083007 (2008).
- [23] J. Chluba and R. A. Sunyaev, *Astron. Astrophys.* **446**, 39 (2006).
- [24] E. E. Kholupenko and A. V. Ivanchik, *Astron. Lett.* **32**, 795 (2006).
- [25] J. Chluba, J. A. Rubino-Martín, and R. A. Sunyaev, *Mon. Not. R. Astron. Soc.* **374**, 1310 (2007).
- [26] E. R. Switzer and C. M. Hirata, *Phys. Rev. D* **77**, 083006 (2008).
- [27] E. R. Switzer and C. M. Hirata, *Phys. Rev. D* **77**, 083008 (2008).
- [28] E. E. Kholupenko, A. V. Ivanchik, and D. A. Varshalovich, *Mon. Not. R. Astron. Soc.* **378**, L39 (2007).
- [29] S. I. Grachev and V. K. Dubrovich, arXiv:0801.3347.
- [30] M. Göppert-Mayer, *Ann. Phys. (Leipzig)* **401**, 273 (1931).
- [31] G. Breit and E. Teller, *Astrophys. J.* **91**, 215 (1940).
- [32] L. J. Spitzer and J. L. Greenstein, *Astrophys. J.* **114**, 407 (1951).
- [33] L. N. Labzowsky, A. V. Shonin, and D. A. Solov'yev, *J. Phys. B* **38**, 265 (2005).
- [34] J. D. Cresser, A. Z. Tang, G. J. Salamo, and F. T. Chan, *Phys. Rev. A* **33**, 1677 (1986).
- [35] V. Florescu, I. Schneider, and I. N. Mihailescu, *Phys. Rev. A* **38**, 2189 (1988).
- [36] J. H. Krolik, *Astrophys. J.* **338**, 594 (1989).
- [37] J. H. Krolik, *Astrophys. J.* **353**, 21 (1990).
- [38] G. B. Rybicki and I. P. dell'Antonio, *Astrophys. J.* **427**, 603 (1994).
- [39] V. B. Berestetskii, E. M. Lifshits, and L. P. Pitaevskii, *Quantum Electrodynamics* (Izdatel'stvo Nauka, Moscow, 1980), 2nd ed..
- [40] J. Chluba and R. A. Sunyaev, *Astron. Astrophys.* **475**, 109 (2007).
- [41] U. D. Jentschura, *J. Phys. A* **40**, F223 (2007).
- [42] U. D. Jentschura, *J. Phys. A* **41**, 155307 (2008).
- [43] G. B. Rybicki, *Escape Probability Methods*, *Methods in Radiative Transfer* (1984), p. 21.
- [44] R. W. Hamming, *Numerical Methods for Scientists and Engineers* (Dover Publications, New York, 1986), 2nd ed..
- [45] U. Seljak and M. Zaldarriaga, *Astrophys. J.* **469**, 437 (1996).
- [46] J. Silk, *Astrophys. J.* **151**, 459 (1968).
- [47] N. Kaiser, *Mon. Not. R. Astron. Soc.* **202**, 1169 (1983).
- [48] W. Hu and N. Sugiyama, *Astrophys. J.* **444**, 489 (1995).
- [49] M. Zaldarriaga and D. D. Harari, *Phys. Rev. D* **52**, 3276 (1995).
- [50] J. Dunkley *et al.*, arXiv:0803.0586.
- [51] G. Hinshaw *et al.*, arXiv:0803.0732.
- [52] M. R. Nolta *et al.*, arXiv:0803.0593.
- [53] R. S. Hill *et al.*, arXiv:0803.0570.
- [54] E. L. Wright *et al.*, arXiv:0803.0577.
- [55] C. L. Reichardt *et al.*, arXiv:0801.1491.
- [56] The Planck Collaboration, arXiv:astro-ph/0604069.
- [57] A. Albrecht *et al.*, arXiv:astro-ph/0609591.
- [58] D. J. Fixsen and J. C. Mather, *Astrophys. J.* **581**, 817 (2002).
- [59] V. K. Dubrovich, *Sov. Astron. Lett.* **1**, 196 (1975).
- [60] G. B. Rybicki and I. P. dell'Antonio, in *Observational Cosmology*, *Astronomical Society of the Pacific Conference Series Vol. 51*, edited by G. L. Chincarini, A. Iovino, T. Maccacaro, and D. Maccagni (Astronomical Society of the Pacific, San Francisco, 1993), p. 548.
- [61] P. Boschan and P. Biltzinger, *Astron. Astrophys.* **336**, 1 (1998).
- [62] J. A. Rubino-Martín, J. Chluba, and R. A. Sunyaev, arXiv:0711.0594.
- [63] B. Podolsky, *Proc. Natl. Acad. Sci. U.S.A.* **14**, 253 (1928).
- [64] A. Maquet, V. Vénier, and T. A. Marian, *J. Phys. B* **31**, 3743 (1998).